Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
COGNITION
Cognition xx (0000) xxx–xxx
www.elsevier.com/locate/COGNIT

Infants chunk object arrays into sets of individuals

Lisa Feigenson*, Justin Halberda

Johns Hopkins University, Baltimore, MD, USA

Received 28 January 2003; revised 4 August 2003; accepted 16 September 2003

Abstract

Research suggests that, using representations from object-based attention, infants can represent only 3 individuals at a time. For example, infants successfully represent 1, 2, or 3 hidden objects, but fail with 4 (*Developmental Science 6* (2003) 568), and a similar limit is seen in adults' tracking of multiple objects (see *Cognitive Psychology 38* (1999) 259). In the present experiments we used a manual search procedure to ask whether infants can overcome this limit of 3 by chunking individuals into sets. Experiments 1 and 2 replicate infants' failure to represent a total of 4 objects. We then show that infants can exceed this limit when items are spatiotemporally grouped into two sets of 2 prior to hiding, leading infants to successfully represent a total of 4 objects. Experiment 3 demonstrates that infants tracked the 4 objects as two sets of 2, searching for each set in its correct hiding location. That infants represented the number of individuals in each set is demonstrated by their reaching for the correct number of objects in each location. These results suggest that by binding individuals into sets, infants can increase their representational capacity. This is the first evidence for chunking abilities in infants.

Keywords: Number; Chunking; Infants

1. Introduction

Psychology has long focused on the limits of what the mind can represent, and the conditions under which those limits can be overcome. A classic example is the study of chunking in short-term memory, in which grouping strongly influences how many items can be represented at any one time (Miller, 1956). While provocative, the origins and mechanisms underlying chunking abilities are not well understood. Here we contribute

* Corresponding author. Department of Psychological and Brain Sciences, 3400 N. Charles Street, Baltimore, MD 21218, USA.
E-mail address: feigenson@jhu.edu (L. Feigenson).

0022-2860/$ - see front matter © 2003 Published by Elsevier B.V.
doi:10.1016/j.cognition.2003.09.003

data that may address these questions by asking whether infants can chunk representations of individual objects into sets.

A foundational finding from the study of memory development is that infant memory is structured much like adult memory (Rovee-Collier, 1999). Pre-linguistic infants have access to an explicit memory system (Adler, Gerhardstein, & Rovee-Collier, 1998), and infants' long-term memory is mediated by the same factors that affect adults' long-term memory (Rovee-Collier, 1999). Infants as young as 3 months old can perform a visual search for a remembered target among distractors (Rovee-Collier, Hankins, & Bhatt, 1992). And, consistent with work on adult visual attention (e.g. Treisman, 1988), infants perform a parallel search when seeking a single feature among distractors (Rovee-Collier et al., 1999) and a serial search when seeking a conjunction of features (Gerhardstein & Rovee-Collier, 2002; Rovee-Collier, Bhatt, & Chazin, 1996).

Work on infant memory has also uncovered the limits of what can be remembered. Beyond simply recognizing targets, infants can remember their serial position when presented in a list (Gulya, Rovee-Collier, Galluccio, & Wilk, 1998; Gulya, Sweeney, & Rovee-Collier, 1999). Also, in an investigation into the limits on the number of individuals infants can store in long-term memory, 3-month-old infants successfully represented information about the features of 2 objects for a 24 h retention period but, given the same exposure time, failed to retain information about 3 objects (Bhatt & Rovee-Collier, 1997a,b). This work has begun to specify the limits of infants' long-term memory. In the present work we ask: might infants, like adults, also have a limit on the number of individuals they can represent in *short*-term memory?

Evidence that infants can represent individuals comes from work on infants' encoding of small arrays of objects (e.g. Strauss & Curtis, 1981; Wynn, 1992). Consider a task from Feigenson, Carey, and Hauser (2002), in which crackers were hidden in two buckets and infants were allowed to crawl to one of them. Infants successfully chose the bucket containing more crackers with comparisons of 1 vs. 2 and 2 vs. 3, showing that they tracked the individuals hidden inside. However, there were limits to infants' abilities. Infants performed at chance with comparisons of 2 vs. 4, 3 vs. 6, and 1 vs. 4, despite the highly discriminable ratios between these numerosities (Feigenson, 2002). This pattern, with infants succeeding with 1, 2, and 3, and failing whenever there were more than 3 individuals in either bucket, shows that infants' performance was limited by the number of individuals per bucket and not by the ratio of the numerosities involved. Since analog magnitude accounts of number predict that performance will not be determined by absolute numerosity, but by the ratio (Weber fraction) between compared numerosities (Gallistel & Gelman, 1992; Whalen, Gallistel, & Gelman, 1990; Wynn, 1998), analog magnitude models do not capture infants' pattern of performance.

Instead, the finding that performance was object-limited with a specific limit of 3 suggests that representations deriving from object-based attention, such as object-files (Kahneman, Treisman, & Gibbs, 1992), underlie infants' encoding of small numbers of objects (see Carey & Xu, 2001; Feigenson et al., 2002; Scholl, 2001; Simon, 1997). Such a claim receives support from empirical demonstrations that adults have a limit on the number of items they can attend to in parallel. This limit is approximately 3–4 (Halberda, Simons, & Wetherhold, 2003; Rensink, 2000; Scholl, Pylyshyn, & Feldman, 2001; Trick & Pylyshyn, 1994; Yantis & Johnson, 1990). Recently, it has also been suggested that the limit of object-based attention and the limit of short-term memory derive from a common source (Cowan, 2001). Irrespective of whether infants have independent limits on attention and on short-term memory, the present work expands on the finding that infants appear limited to representing approximately 3 items at any one time.

Tasks such as the one by Feigenson et al. show that infants can represent multiple individuals and store them in short-term memory. Infants can also perform computations over these representations of individuals. In Feigenson et al.'s choice task, infants were able to sum overall cracker volume to determine which bucket contained more, choosing 1 large cracker over 2 small, and performing at chance when volume was equated between the choices (Feigenson et al., 2002). Additionally, infants can compute one-to-one correspondence between object-files, a computation equivalent to assessing whether two arrays contain the same number of objects (Feigenson & Carey, 2003).

The pattern of performance in the above tasks motivates a question about how many individuals infants can represent at one time. Given that infants in the Feigenson et al. choice task succeeded at comparing 2 vs. 3 crackers, perhaps the 3-item representational limit is not a global one. Feigenson et al. (2002) proposed that infants are limited to representing up to 3 individuals *per bucket*. Thus, infants in the choice task can represent 2 crackers in one bucket and 3 in another. But once the number in either bucket exceeds 3, the representation falls apart. Crucially, infants failed at a 1 vs. 4 comparison (Feigenson, 2002) but succeeded with 2 vs. 3. This suggests the possibility that by grouping individuals into distinct sets, each of 3 or fewer, infants might be able to represent arrays that would otherwise exceed their capacities. For example, infants might fail to represent a single array of 5 individuals, but succeed at representing a set of 2 and a set of 3 because there are 3 or fewer individuals per set.

In this sense, the binding of objects into sets of objects is analogous to the classic phenomenon of chunking by adults. In some cases, such as that of the famous S.F. who increased his digit span by chunking individual numerals into race times, the individuals that make up a chunk play a critical role in defining the chunk, and can be retrieved as individuals later on (Ericsson, Chase, & Faloon, 1980). This raises the question whether infants, like adults, can form a representation of a set wherein they maintain both a representation of the set *and* a representation of the individuals that form the set.

The data from Feigenson et al. (2002) are suggestive, but they do not directly answer this question because success on the cracker-task does not require infants to represent more than 3 *individuals* at one time. Although the degree to which infants systematically chose one bucket over the other was determined by the *number* of crackers presented, which bucket infants chose was determined by the total *amount* of crackers in each bucket. For example, infants chose 1 large over 2 small crackers (Feigenson et al., 2002). Feigenson et al. suggested that infants created object-file representations of each individual cracker as it was presented, then summed across those representations to achieve an estimate of the "total cracker material" per bucket. Though infants were limited by the number of available object-files to encoding 3 crackers per bucket, infants' choice was governed by how much "cracker-material" was in each bucket irrespective of the number of individuals. Because this representation collapses over individuals, infants' success at choosing between 2 vs. 3 crackers does not provide evidence that infants can represent more than 3 individuals in parallel or that infants can chunk representations of individuals into sets.

Two other representations that collapse across individuals, and thus do not show evidence of set-building by infants, are analog magnitudes and perceptual groups. It is well documented that infants can represent approximate number information for groups of objects at least as numerous as 32 dots (Xu & Spelke, 2000), and for sequences of sounds at least as numerous as 16 beeps (Lipton & Spelke, 2003). That these infants are using the approximate number system of analog magnitudes is suggested by the fact that performance varies with the Weber fraction of the numerosities presented (Lipton & Spelke, 2003; Xu & Spelke, 2000). But because analog magnitudes are holistic representations, infants' success in discriminating, for example, 16 from 32 dots does not show that they have overcome the 3-item limit of parallel attention. When an infant represents "approximately 32", the infant is entertaining a single representation (Gallistel & Gelman, 2000). Empirical data as well as the current models of how infants form analog magnitudes suggest that infants do so without ever applying focal attention to the individuals that comprise the collection (Barth, Kanwisher, & Spelke, 2003; Church & Broadbent, 1990; Dehaene & Changeux, 1993). Clearly then, analog magnitudes do not show evidence that infants can represent *individuals* in parallel attention and chunk these individuals into sets.

For this same reason, infants' ability to match numerosity across modalities does not show that infants have exceeded the 3-item limit of parallel attention. In inter-modal matching studies, infants habituated to 3 sounds will increase visual attention to a display containing 3, rather than 2, dots (Feron, Streri, & Gentaz, 2002; Kobayashi, Hiraki, & Hasegawa, 2002; Starkey, Spelke, & Gelman, 1990). However, as described above, the current evidence suggests that infants represent the numerosity of a sequence of sounds via an analog magnitude representation. Thus, infants may represent the abstract numeric similarity between sounds and dots without applying attention to and storing a representation of *each individual* in the array (Gallistel & Gelman, 1992). And, these two limits, i.e. visual attention and auditory attention, have been shown to be independent in adults (Scholl & Xu, 2001). Thus, though infants may represent 6 dots and 6 sounds simultaneously, this in no way demonstrates that they have overcome the representational limit of 3 individuals in parallel attention.

Besides the ability to represent large numerosities in analog form, infants also have the ability to represent perceptual groups of objects. Wynn, Bloom, and Chiang (2002) showed that 5-month-old infants respond to a change in the number of collections in an array, where each collection was comprised of multiple dots moving across the screen together. In their study, infants in one condition were habituated to 2 collections of 3 dots each (total = 6 dots). They were then tested with 2 collections of 4 dots, vs. 4 collections of 2 (total = 8 dots for both test types). Infants looked longer at the array with the novel number of collections. This result is important because it demonstrates that infants can enumerate entities that are composed of smaller objects, treating each collection as a perceptual group. However, it does not show evidence of the kind of chunking abilities we seek because it does not show that infants represented both the set (the collection) and the individuals comprising the set (the dots). In order to succeed at Wynn et al.'s task, infants need only pay attention to the number of low-level perceptual groups in the array.

These results are important ingredients to studying chunking abilities in infants. Feigenson et al. show that the 3-object limit of parallel attention may not be a global one, and Wynn et al. show that when provided with appropriate spatiotemporal information, infants can treat multiple individuals as a single perceptual group. In the present paper we seek evidence that infants can chunk individuals into sets. This capacity requires two levels of representation, the *set* and the *individual*, and must allow the retention of information about individuals even after the individuals have been bound (or chunked) into a set. We hypothesize that such set-building can allow infants to exceed the 3-item limit of parallel attention.

We test this hypothesis in three experiments. Experiment 1 asks whether infants can use spatiotemporal information to represent two sets of 2, thereby representing a total of 4 individuals. Experiment 2 replicates the findings from Experiment 1 and manipulates the availability of set-building information as a within-subject factor. Experiment 3 asks whether infants can track two sets as they move to independent spatial locations. In all of these experiments, infants are required to represent the total number of individuals, and not just the number of sets, in order to succeed.

2. Experiment 1

Infants' failure to represent 4 total individuals in Feigenson and Carey's (2003) task serves as a basis of comparison and a procedural model for the present series of experiments. In their study, 14-month-old infants saw an experimenter present an array of identical balls on top of a box. The experimenter then hid the balls inside and allowed infants to reach in and retrieve either all of them, or just a subset. Infants' subsequent searching in the box provided a measure of how many objects they represented inside.

Crucially, infants in this task have been shown to respond based on the *number* of objects presented rather than on analog magnitudes or overall amounts of material. Infants who saw 2 small objects hidden in a box continued searching after they had retrieved 1 *large* object, showing that their searching was guided by a representation of how many objects were in the box, and not by how much object-material was in the box (Feigenson & Carey, 2003).

Feigenson and Carey presented infants with numerical comparisons of 1 vs. 2, 2 vs. 3, and 2 vs. 4. Each of these pairs contained a measure of infants' searching when the box was expected to be empty, and when it was expected to contain more balls. For example, a 2 vs. 4 comparison compared searching when infants saw 2 balls hidden and had retrieved 2 (Box Empty) with searching when infants saw 4 balls hidden and had retrieved only 2 (More Remaining). Subtracting search times on Box Empty measurement periods from those on More Remaining measurement periods creates a difference score. If infants accurately represent the total number of balls in the box, they should search more on More Remaining than on Box Empty measurement periods, revealing positive difference scores.

The results of the 1 vs. 2 and 2 vs. 4 comparisons are re-printed in Fig. 2a as difference scores, and in Fig. 3a by individual measurement period. While infants successfully made numerical discriminations in the 1 vs. 2 condition (by searching the box after seeing 2 hidden and retrieving just 1), they failed with 2 vs. 4. This can be seen as a positive difference score on 1 vs. 2 comparisons, whereas with the 2 vs. 4 comparisons the difference score was not different from chance.

Thus, 14-month-old infants succeeded at representing 1, 2, and 3[1] hidden individuals but failed to represent 4. This is striking because longer searching on 4-Object (More Remaining) measurement periods does not require an exact representation of 4, merely that 4 is more than 2. This evidence converges with Feigenson et al. (2002) in demonstrating an abrupt limit of 3 on infants' ability to represent individuals.

If infants' failure in the 2 vs. 4 comparison of Feigenson and Carey's task is explained by their inability to represent a single set of 4 individuals, can infants successfully represent two sets of 2? In contrast to Feigenson and Carey (2003), who presented all of the balls in a single set atop the box, we provided infants with spatiotemporal information to help them group individuals into two distinct sets before we hid them. This can be seen by comparing Fig. 2a, which depicts the key 2- and 4-object presentations of Feigenson and Carey, and Fig. 2b, which depicts the 2- and 4-object presentations in Experiment 1. All other aspects of the two procedures were identical.

2.1. Method

2.1.1. Participants

Sixteen 14.5-month-old infants participated (mean = 14 months, 14 days). Ten were boys. Two additional infants were excluded due to fussiness (1) or failure to search (1).

2.1.2. Stimuli

Infants watched the experimenter hide ping-pong balls in a foam-core box ($31.5 \times 25 \times 12.5$ cm). The box's face had a 14×7.5 cm opening covered by cloth with a slit across its width, and a felt-covered opening at the rear. Two 12×12 cm foam-core platforms rested 12 cm from either side of the box.

2.1.3. Procedure

Infants sat in a high chair in front of a table, with the experimenter kneeling to the left. A video camera recorded a side-view of the session. Infants received one block each of 1 vs. 2 and 2 vs. 4 comparisons.

2.1.3.1. 1 vs. 2 comparisons. Fig. 1 shows the presentation time-course of 1 vs. 2 comparisons. One-Object (Box Empty) measurement periods measured searching after infants saw 1 ball hidden and had retrieved it. First, the box was placed on the table, out of the infants' reach. Then the experimenter brought out 1 ball from a hidden cache of balls and set it on one of the platforms beside the box. She pointed and said, "Look at this." Then, to equate motion and presentation length with those in the 2-Object trials, she pointed to the empty space on the other platform and said, "Look at this." Finally, she picked up the ball, inserted it through the box's opening, pushed the box forward, and said,

[1] Although not shown here, infants in Feigenson and Carey's (2003) task also succeeded with a 2 vs. 3 comparison.

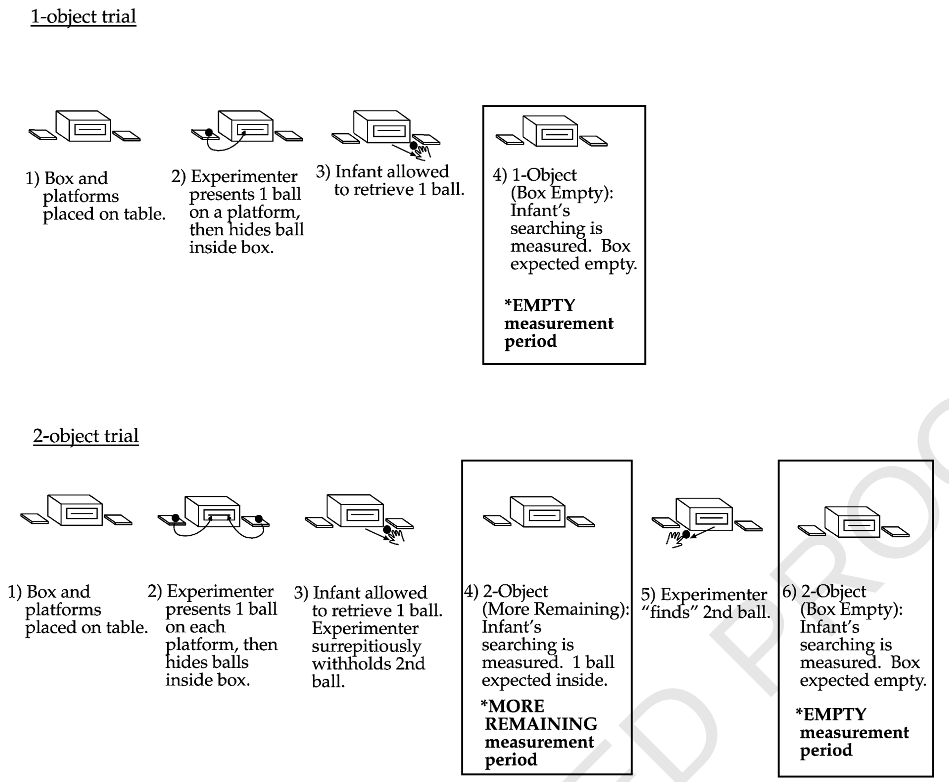

Fig. 1. vs. 2 comparisons in Experiment 1. Difference scores were obtained by subtracting the average of searching time on 1-Object (Box Empty) and 2-Object (Box Empty) trials from searching time on 2-Object (More Remaining) trials. Positive difference scores indicate successful discrimination of 1 vs. 2 objects.

"What's in my box?" Infants were then allowed to retrieve the ball and play with it for several seconds before the experimenter took it away.

After the ball was removed from the infants' hands, a silent 10 s measurement period followed in which the box was left in place and any searching was coded later from videotape. During this period, the experimenter kept her head down and did not look at the infant in order to avoid providing any cues. For a behavior to count as searching, one or both of the infants' hands had to be inserted into the box up to the third knuckle. The measurement period always began immediately after the experimenter took the just-retrieved ball away from the infant, and was not dependent on when the infant actually reached into the box. Indeed, on some measurement periods infants did not reach at all. After 10 s, the experimenter removed the box and the trial ended. If infants were in the middle of searching at the end of 10 s, the experimenter allowed that reach to terminate before removing the box.

Two-Object (More Remaining) measurement periods measured searching after infants saw 2 balls hidden and had retrieved only 1. The experimenter brought out 2 balls from the cache and set one on the right platform and the other on the left. Identical with the procedure for 1-Object trials, she pointed to each and said, “Look at this”, then picked up both balls in one hand and inserted them through the box’s opening. As she inserted the balls in the box, she surreptitiously moved one of the balls to the very back of the box, holding it out of reach but still inside. Hence, infants saw 2 balls hidden, but could only retrieve 1. The box was made large enough so that infants’ hands could never touch the 2nd ball that was being withheld by the experimenter. Furthermore, to ensure that all movements were identical in the 1- and 2-Object trials, the experimenter’s hand remained inside the rear of the box on all trials.

As in the 1-Object (Box Empty) measurement period, infants were allowed to retrieve 1 ball and handle it before it was taken away. A silent 10 s measurement period followed, identical to the 1-Object (Box Empty) measurement period. Because the experimenter had moved the 2nd ball out of the infants’ reach, no evidence of a second ball was present during a 2-Object (More Remaining) measurement period that was not present during a 1-Object (Box Empty) measurement period. If infants can successfully represent 2 objects in the box, they should search for the 1 object still expected inside.

After 10 s, the experimenter reached into the front of the box and “retrieved” the 2nd ball. She gave it to infants and allowed them to handle it briefly. Once it was taken away the last 10 s silent measurement period began, during which the experimenter kept one hand in the back of the empty box to ensure that all three measurement periods were identical. This was called the 2-Object (Box Empty) measurement period because infants had seen 2 balls hidden, had retrieved both, and now the box was empty again. If infants correctly represented 2 objects in the box, searching should return to its baseline Box Empty rate. After 10 s, the trial ended and the experimenter removed the box. Infants received two presentations of each of these three trial types. Whether the 1-Object or the 2-Object trial was presented first was counterbalanced. Two-Object (Box Empty) measurement periods always occurred after 2-Object (More Remaining) measurement periods. The overall pattern of searching that would indicate successful discrimination of 1 vs. 2 objects is: little searching on the 1-Object (Box Empty) measurement period, more searching on the 2-Object (More Remaining) measurement period, and little searching again on the 2-Object (Box Empty) measurement period.

2.1.3.2. 2 vs. 4 comparisons. These trials were structured identically to those in the 1 vs. 2 comparisons. The experimenter either placed 1 ball on each platform (in the 2-Object presentation), or 2 balls on each platform (in the 4-Object presentation, seen in Fig. 2b). As with the 1 vs. 2 comparisons, the amount of motion, length of exposure, and the verbal and non-verbal attention drawn to the balls was identical for the 2-Object and 4-Object trials. And as before, during the 10 s measurement periods the experimenter always maintained a downward gaze to avoid the possibility of cueing infants.

The 2-Object (Box Empty) measurement period measured searching after infants saw 2 balls hidden and had retrieved both. The 4-Object (More Remaining) measurement period measured searching after infants saw 4 balls hidden and had retrieved only 2. As in Feigenson and Carey (2003), the 4 balls were always placed into the box 2 at a time. Here, as with the 1 vs. 2 comparisons, the experimenter surreptitiously withheld the “extra” balls in the rear of the box. For 2 vs. 4 comparisons, this involved holding 2 of the 4 balls against the felt-covered opening in the rear of the box so that they were out of the infants’ range of

COGNIT 1277—28/10/2003—12:04—HEMAMALINI—86226 – MODEL 1a

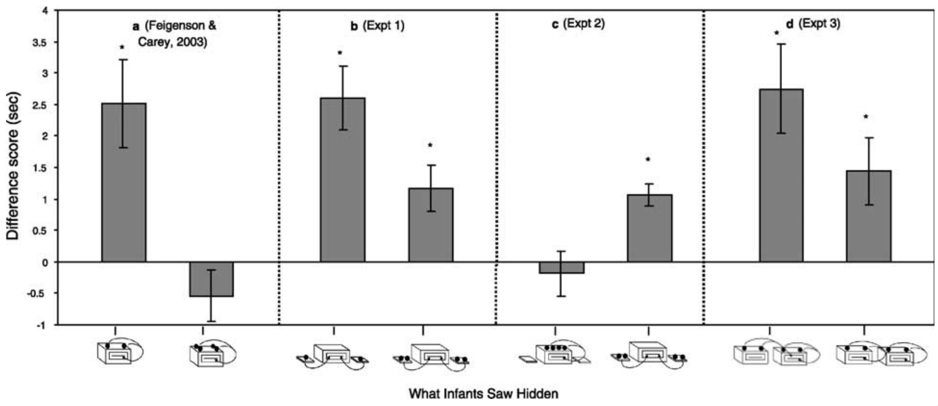

Fig. 2. Infants' performance is displayed as a series of difference scores, computed as searching on More Remaining trials minus searching on Box Empty trials. The *x*-axis displays the number and presentation of balls in each condition, with only the greater number of balls in each comparison displayed. Therefore, bars depict success or failure at representing the number of objects shown in each corresponding schematic. (a) In Feigenson and Carey (2003), infants succeeded with a 1 vs. 2 comparison, but failed with 2 vs. 4. (b) In Experiment 1, infants succeeded with 1 vs. 2 and 2 vs. 4 when balls were presented in spatiotemporally defined sets. (c) In Experiment 2, infants succeeded with 2 vs. 4 only when balls were presented in sets. (d) In Experiment 3, infants again succeeded with 2 vs. 4, and tracked the separate hiding locations of the two sets.

reach. One of the experimenter's hands remained in the rear of the box on all trials in order to ensure that 2-Object and 4-Object trials were identical in all respects. Before the 10 s measurement period could begin, infants were required to retrieve 2 balls, i.e. infants reached in and retrieved 1 ball, the experimenter took it away, and infants reached again and retrieved a 2nd ball. Fourteen of the 16 infants tested spontaneously reached in twice and retrieved both balls. The remaining two infants required help from the experimenter in retrieving the 2nd ball on one out of the four 4-Object trials.

Finally, the 4-Object (Box Empty) measurement period measured searching after infants were given the remaining 2 balls from the 4-Object (More Remaining) presentation.

In sum, the only difference between the methods of Experiment 1 and Feigenson and Carey (2003) was the presentation of the balls. Here, instead of placing them atop the box, the experimenter always placed them in two sets on the side platforms.

Searching was coded from videotape by two observers. Agreement between the two observers was 95%.

2.2. Results and discussion

Unlike in Feigenson and Carey (2003), infants in Experiment 1 succeeded at discriminating both the 1 vs. 2 and 2 vs. 4 comparisons (presented as difference scores in Fig. 2b, and by measurement period in Fig. 3b). A 2 (Block: 1 vs. 2 or 2 vs. 4) × 2 (Block Order) × 4 (Trial Order: within a block, whether the larger number was presented first or second) × 3 (Measurement Period: first Empty period, More Remaining period, second Empty period) ANOVA revealed a main effect of Measurement Period ($F(2, 16) = 21.77$,

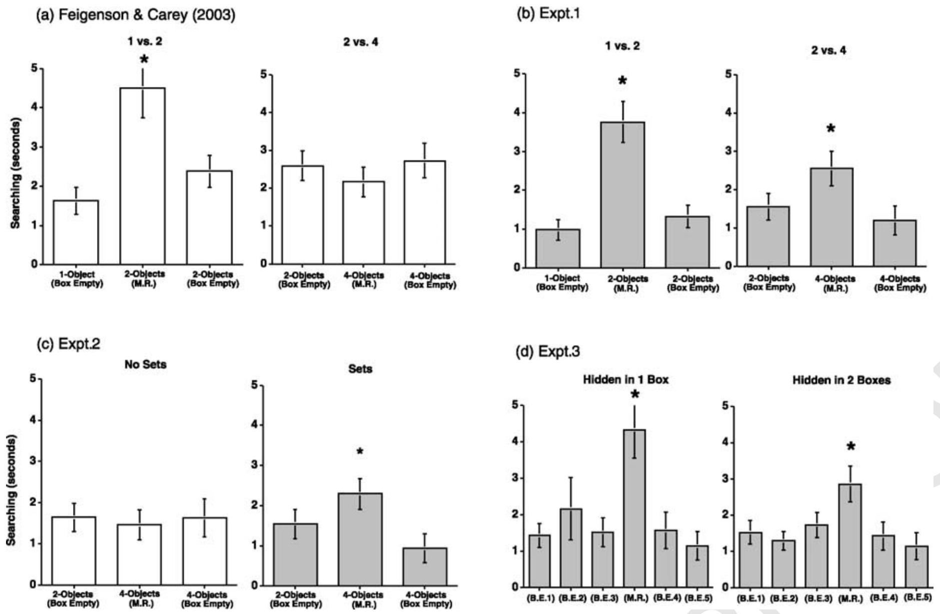

Fig. 3. Infants' performance by measurement period. MR denotes More Remaining periods, and BE denotes Box Empty periods. Dark bars represent arrays presented in spatiotemporally distinct sets. Open bars represent arrays not presented in sets. In (a), infants successfully differentiated 1 vs. 2 hidden objects but not 2 vs. 4. In (b), infants differentiated both 1 vs. 2 and 2 vs. 4 when objects were presented in two sets. In (c), infants differentiated 2 vs. 4 only on trials when objects were presented in sets. In (d), infants tracked the location of each set of 2 objects, regardless of their hiding location.

$P < 0.001$), and a Block × Measurement Period interaction ($F(2, 16) = 5.28$, $P < 0.05$). That is, infants searched differentially by Measurement Period, and succeeded more strongly in the 1 vs. 2 than the 2 vs. 4 comparison. We investigated the source of these effects with planned *t*-tests.

In the 1 vs. 2 block the two types of Box Empty measurement periods did not differ ($t(1, 15) = -1.27$, $P = 0.22$) (Fig. 2b), and so were collapsed into one Box Empty search time. Subtracting this from More Remaining searching created a difference score of +2.35, which differed from the difference score of zero predicted by chance searching ($t(1, 15) = -4.68$, $P < 0.001$) (Fig. 2b). In the 2 vs. 4 block the two types of Box Empty measurement periods also did not differ ($t(1, 15) = -0.84$, $P = 0.41$) (Fig. 3b), and so were also collapsed. Subtracting this from More Remaining searching yielded a difference score of +1.17, which also differed from chance ($t(1, 15) = -3.18$, $P < 0.01$) (Fig. 2b). Thus, for both the 1 vs. 2 and 2 vs. 4 comparisons, infants searched more on More Remaining than on Box Empty measurement periods.

Infants in both Experiment 1 and in Feigenson and Carey (2003) succeeded with the 1 vs. 2 comparison. But only infants in the present experiment successfully discriminated 2 vs. 4. All aspects of the experiments were identical, including all movements of the balls, presentation times, and experimenter verbalizations. The only exception was that the balls were presented as a single set of 4 in Feigenson and Carey's experiment, and as two sets of

2 here in Experiment 1. This suggests that spatiotemporal grouping determined whether infants could represent the total number of individuals. Note that infants could not have succeeded on 2 vs. 4 comparisons simply by representing the number of *sets* presented (i.e. 2), or by representing these as low-level perceptual groups. On 2 vs. 4 comparisons infants reached into the box and retrieved 2 balls even before the "More Remaining" measure began. Therefore, infants' increased "More Remaining" searching was always comprised of their third and fourth reaches into the box, demonstrating that they represented not only the number of sets, but also the correct total number of hidden objects.

3. Experiment 2

Infants in Experiment 1 succeeded with a 2 vs. 4 comparison. In contrast, the 14-month-old infants in Feigenson and Carey's (2003) study failed with a 2 vs. 4 comparison when the 4 balls were placed in a square on top of the box. But even in their study, the square configuration provided some evidence that could have led infants to represent these 4 balls as two sets of 2. And, consistent with the presentation in Experiment 1, Feigenson and Carey placed the 4 balls on top of the box two at a time, and hid them in the box two at a time. Therefore, the *only* difference between these two procedures was the presentation of the balls on the platforms in Experiment 1. It is surprising that such a small change led infants to overcome the 3-item limit on parallel attention. Thus, Experiment 2 sought to replicate the results of Experiment 1 under even more stringent conditions. Infants received only 2 vs. 4 trials. On half of them, infants were given spatiotemporal information for sets as in Experiment 1: 4 balls were presented as two sets of 2 on the separated platforms (Two Sets block, Fig. 2c). On the other half, 4 balls were presented in a single line on top of the box (Single Set block, Fig. 2c). All of the movements in the presentation and hiding of the balls were identical between the two types of trials. Thus, we sought within-subject evidence that spatiotemporal grouping can determine how many individuals infants can represent.

3.1. Method

3.1.1. Participants

A new group of 16 14.5-month-old infants participated in Experiment 2 (mean = 14 months, 15 days). Five were boys. One additional infant was excluded due to fussiness.

3.1.2. Procedure

The design and procedure were identical to those in Feigenson and Carey (2003) and Experiment 1, except that infants received only 2 vs. 4 comparisons. All movements, timing, length of exposure, and the verbal and non-verbal attention drawn to the balls were identical for these 2 vs. 4 comparisons, with the sole difference being the presentation locations of the balls. They were either presented on the side platforms as two sets of 2 (Two Sets), or in a single line on top of the box (Single Set).

As in Experiment 1, searching was coded from videotape by two observers. Agreement between the two observers was 94%. Furthermore, Experiment 2 took two additional measures to ensure that there was no possible bias in the experimental presentation that could have influenced infants' searching patterns. First, a pair of observers measured from videotape the amount of time the experimenter looked at the infant during each 10 s measurement period, during which the experimenter had been instructed to look down to avoid providing the infant with any cues. These observers found that the experimenter looked at the infant only 5% of the total measurement time on Single Set trials, and only 6% of the time on Two Sets trials. Therefore, infants could not have differentiated trials in Single Set blocks (in which we predicted failure to differentiate 2 vs. 4) from those in Two Sets blocks (in which we predicted successful differentiation) based on the amount of time the experimenter was looking at them or at the box.

A second measure ascertained whether the experimenter had cued the infants in any other respect. Two observers watched each 10 s measurement period on video and tried to guess what type of measurement period it was, using all available auditory and visual information. Videotapes were cued by a third person so that the two observers were completely blind as to whether the measurement period was in a Single Set or a Two Sets block, as to how many balls had been hidden, and as to how many balls (if any) had already been retrieved by the infant. Both observers performed at chance levels for guessing the type of measurement period. Crucially, on the 4-Object (More Remaining) periods, the observers were at chance at guessing whether the balls had been presented in a single set or in two sets (Observer 1: 54% correct; Observer 2: 43% correct). These results mitigate against any possible presentation bias on the critical measurement periods.

3.2. Results and discussion

In Experiment 2, infants' success depended on whether the balls were presented in two sets or in a single set. A 2 (Block: Single Set or Two Sets) × 2 (Block Order) × 3 (Measurement Period: first Empty period, More Remaining period, second Empty period) ANOVA found only a Block × Measurement Period interaction ($F(2,28) = 3.61$, $P < 0.05$). That is, infants differentiated the measurement periods only in the Two Sets block. We investigated this effect with planned *t*-tests.

In the Single Set block the two types of Box Empty measurement periods did not differ ($t(1,15) = 0.01$, $P = 0.99$) (Fig. 3c), and so were collapsed into a single Box Empty search time. Subtracting this from More Remaining searching created a difference score of -0.13, which did not differ from the difference score of zero predicted by chance ($t(1,15) = 0.34$, $P = 0.74$) (Fig. 2c). In the Two Sets block the two types of Box Empty measurement periods also did not differ ($t(1,15) = 1.56$, $P = 0.14$) (Fig. 3c), and so were also collapsed. Subtracting this from More Remaining searching yielded a difference score of $+1.06$, which was different from chance ($t(1,15) = -6.4$, $P < 0.001$) (Fig. 2c). That is, only when the balls were presented as two spatiotemporally defined sets did infants reach more for 4 balls than for 2.

These results replicate Feigenson and Carey's finding that infants fail with 2 vs. 4 when 4 objects are presented in a single set. They also strengthen the results of Experiment 1 by showing that even when treated as a within-subject variable, the presentation of the balls as two sets of 2 determined whether infants could track 4 total individuals.

4. Experiment 3

Are infants in fact tracking the 4 objects as two sets of 2? In Experiments 1 and 2, because all 4 objects were eventually hidden in the same box, it is possible that the spatial grouping of the objects prior to hiding simply made it easier for infants to represent and remember 4 separate balls, and not two sets of 2. If we are correct in claiming that infants are representing two distinct sets, they should be able to track these sets even if they move to separate locations. Experiment 3 tested this by adding a second box in which the balls could be hidden. We again used only 2 vs. 4 comparisons. On one block, all of the balls were hidden in one of the boxes. On the other block, 2 of the balls were hidden in one box, and 2 in the other (Fig. 2d). If infants are in fact tracking the 4 balls as two sets of 2, they should succeed at both "One Box" and "Two Box" comparisons.

4.1. Method

4.1.1. Participants

A new group of 16 14-month-old infants participated in Experiment 3 (mean = 13 months, 21 days). Ten were boys. Five additional infants were excluded due to fussiness.

4.1.2. Procedure

Experiment 3 used a procedure similar to that of Experiment 2, but with the addition of a second box (Fig. 2d). Both boxes were on the table during all trials. These boxes served as the separate presentation locations for the balls throughout the study, analogous to the platforms in Experiments 1 and 2. On 2-Object trials, 1 ball was placed atop the right-hand box, and 1 atop the left-hand box. On 4-Object trials, 2 balls were placed atop each box using the same timing and motion controls employed throughout Experiments 1 and 2. Infants received a One Box block and a Two Box block of trials. On One Box trials, all of the balls (either 2 or 4) were hidden in one of the two boxes. On Two Box trials, the balls on top of each box were hidden inside that box, thereby sending the sets to two separate hiding locations (Fig. 2d).

The inclusion of the second box added 3 more Box Empty measurement periods to each 2 vs. 4 comparison. These arose because infants could always reach into either of the two boxes on any given trial. The structure of the trials, including descriptions of all of the measurement periods, is described in Table 1 (One Box block) and Table 2 (Two Box block).

4.2. Results and discussion

Infants successfully discriminated 2 from 4 on both One Box and Two Box blocks. A 2 (Block: One or Two Boxes) × 2 (Block Order) × 4 (Trial Order) × 6 (Measurement Period) ANOVA yielded a main effect of Measurement Period ($F(5,40) = 9.57$,

Table 1
The structure of the trials in the One Box block of Experiment 3, with time extending from left to right

Trial type	Hiding of balls	Infant retrieves	Measurement period
2-Object One-Box	2 balls hidden in Box A; Box B is empty[a]	2 balls from Box A[b]	Box Empty 1: measures any further searching in Box A Box Empty 2: measures any searching in Box B
4-Object One-Box	4 balls hidden in Box A; Box B is empty	2 balls from Box A (2 are secretly being withheld)[c] After 10 s, experimenter retrieves remaining 2 balls	Box Empty 3: measures any searching in Box B More Remaining: measures any further searching in Box A Box Empty 4: measures any further searching in Box A Box Empty 5: measures any further searching in Box B

[a] The label Box A does not refer to a unique box constant across all trials, but is rather used to guide the reader through an individual trial. Box A was whichever box the balls were hidden in (counterbalanced for being the box on the left vs. right). Box B simply refers to the other box, in which no balls were hidden.
[b] On 91% of all 2-Object trials infants retrieved both balls without any assistance from the experimenter.
[c] On 88% of all 4-Object trials infants retrieved both balls without any assistance from the experimenter.

$P < 0.001$), showing that infants searched differentially by trial type. We investigated this effect with planned *t*-tests.

In the One Box block, the 5 Box Empty measurement periods did not differ from each other ($F(4, 60) = 0.78$, $P = 0.54$) (Fig. 3d), and so were collapsed into one Box Empty

Table 2
The structure of the trials in the Two Box block of Experiment 3, with time extending from left to right

Trial type	Hiding of balls	Infant retrieves	Measurement period
2-Object Two-Box	1 ball hidden in Box A; 1 ball hidden in Box B	1 ball from Box A and 1 ball from Box B[a]	Box Empty 1: measures any further searching in Box A Box Empty 2: measures any searching in Box B
4-Object Two-Box	2 balls hidden in Box A; 2 balls hidden in Box B	Either 1 ball from Box A and 1 ball from Box B, or 2 balls from either Box A or Box B[b] After 10 s, experimenter retrieves remaining 2 balls	Box Empty 3: if infants retrieved 2 balls from the same box, measures any further searching in that box More Remaining: if infants retrieved 1 ball from each box, measures any further searching in either box; if infants retrieved 2 balls from the same box, measures any further searching in the other box Box Empty 4: measures any further searching in Box A Box Empty 5: measures any further searching in Box B

[a] On 84% of all 2-Object trials infants retrieved both balls without any assistance from the experimenter.
[b] On 90% of all 4-Object trials infants retrieved both balls without any assistance from the experimenter.

search time. Subtracting this from More Remaining searching created a difference score of +5.51, which differed from chance ($t(1,15) = 3.92$, $P < 0.01$) (Fig. 2d). In the Two Box block the 5 Box Empty measurement periods also did not differ ($F(4,60) = 0.45$, $P = 0.77$) (Fig. 3d), and so were also collapsed. Subtracting this from More Remaining searching yielded a difference score of +2.87, which again differed from chance ($t(1,15) = 2.66$, $P < 0.05$) (Fig. 2d). Thus, infants differentiated those measurement periods in which more balls remained to be found from all other measurement periods in both blocks.

Infants succeeded with 2 vs. 4 comparisons whether balls were hidden in one or two boxes. Because searching in the wrong box (even when more balls were expected in the correct box) counted as Box Empty searching, infants' increased searching on More Remaining measurement periods indicates that they knew where each set of 2 balls was located and, because infants had to reach at least two times to retrieve each of the 2 balls in a set, infants' reaching also reveals that they also knew how many balls each set contained.

5. General discussion

All chunking abilities share a common underlying cognitive architecture: the ability to maintain at least two levels of representation. These two levels are that of the individual and that of the set or chunk. Here we see the possible cognitive origins of this ability: infants' ability to maintain a representation of two sets of objects and the individuals that make up each set.

In previous reports (Feigenson & Carey, 2003; Feigenson et al., 2002), infants presented with small numbers of objects have shown a representational limit of 3 individuals, coincident with the limit on adults' performance in visual search and multiple object tracking tasks (see Rensink, 2000; Scholl & Pylyshyn, 1999). In Experiments 1–3 we show that this limit can be exceeded when objects are spatiotemporally grouped to form smaller sets. Infants successfully searched for two sets of 2 objects in the appropriate locations, while failing with a single set of 4.

How do these results bear on classic examples of chunking in short-term memory? A multiple-level cognitive architecture that allows for an increase in representational capacity is shared by every example of chunking. The hierarchical encoding of individuals and sets of individuals revealed in the present experiments is one example, similar to that demonstrated by the subject S.F. (Ericsson et al., 1980), who vastly increased his digit-span memory using the conscious strategy of chunking numerals into race times. S.F. stored 4 5 1 1 not as 4 separate numerals ([4], [5], [1], [1]), but as a single race time of 45.11 s ([4,5,1,1]). By creating a hierarchy of such levels, S.F. was able to increase his digit span from 7 to 80 digits. We do not suggest that infants are using such a conscious strategy, but rather that the spatiotemporal evidence for two distinct locations of objects, and the shift in attention required as the second set of objects is presented, motivates a second level of encoding.

How might this ability to bind representations of objects into sets develop into more sophisticated chunking abilities? Beyond the power of recursion noted in the example of S.F., S.F.'s abilities evidence another feature of classic chunking: the ability to use diverse sources of information (in his case, stored semantic content about race times) to motivate the formation of chunks. In the present studies, infants bound individuals into sets based on Gestalt grouping cues such as proximity (Wertheimer, 1923) and common region (Palmer & Rock, 1994). Again, evidence that infants are doing more than merely perceptually grouping these individuals comes from the finding that infants matched their searches to the number of total individuals hidden. Thus, while the basis of the set-building was perceptual rather than semantic, infants' performance shows the key features of overcoming a representational limit and maintaining access to the individuals comprising the set.

Might infants, like adults, also be able to bind objects into sets using other more diverse sources of grouping information? Leslie has obtained evidence that this may be the case: he argues that 12-month-old infants form the representation "pair of objects", where a pair is defined by common shape (e.g. triangles vs. discs) or common color (e.g. red vs. yellow) (Leslie, 2003; Leslie & Glanville, 2002). In these experiments, infants familiarized to an array containing (triangle, triangle, disc, disc) looked longer at test outcomes of (triangle, disc, triangle, disc) than at outcomes that preserved the original configuration. If infants formed the representation "pair", where "pair" implies the presence of precisely 2 objects, then this would be evidence that infants can form sets based on property information. However, it remains an open question whether this is the case. Leslie did not present infants with outcomes that were numerically novel (such as (triangle, triangle, triangle, disc, disc, disc)). Therefore, we cannot know if infants actually represented the correct number of individuals in the array, or if they were responding to a change in a global property such as the overall symmetry of the array (i.e. representing "pair" as something like "same kind of thing" with no numerical commitment).

If infants can bind representations of individuals into sets based on property information or semantic information, this would more closely align the abilities demonstrated here with the classic work on adult chunking. Current studies in our lab are addressing this question by asking whether infants can form sets based on object-kind information. In the absence of strong spatial-grouping cues, we present infants with 2 animals and 2 artifacts and ask how many total objects they can represent. If successful, this work will continue to form a bridge between early set-building abilities and the more sophisticated abilities seen in adults.

Finally, in order to emphasize the distinction between the set-building computations described here and other types of grouping abilities, we propose that examples of set-building must meet three criteria. First, they must demonstrate that participants are attending to *individuals* and that there is a limit on the number of individuals that can be represented in parallel (see Cowan, 2001 for a review of the evidence that this limit is approximately 3 across many tasks using many types of stimuli). Second, they must show that this limit can be exceeded by chunking individuals into sets of individuals. Third, they must demonstrate that the individuals comprising the chunk or set can be recovered from memory. Knowing that a phone number is comprised of two chunks is of little use if we are unable to recover the individual digits that make up the chunks. The present studies meet these three criteria.

Future studies will continue to explore the dimensions that can motivate the formation of a set, and will also investigate the limits on infants' set-building abilities by asking how many sets infants can represent and how many individuals can be represented within a set. Predictions can be made from the existing data. We predict that infants will be able to represent no more than 3 sets consisting of no more than 3 individuals each. This is because the 3-item limit on parallel attention should prevent infants from simultaneously representing more than 3 objects, a condition necessary for binding the representations into a set. Such investigations will more clearly define how the early set-building abilities presented here add to the growing intersection of the domains of visual attention, short-term memory, and numerical cognition.

6. Uncited References

Simon and Chase, 1973. Spelke, 1994. Uller et al., 1999. Xu et al., 2003.

Acknowledgements

This work was supported by National Institutes of Health (NIH) grant HD-38338-01 to Susan Carey, and National Science Foundation (NSF) predoctoral fellowships to L.F. and J.H. The authors gratefully thank Susan Carey, who provided inspiring discussion throughout this project, as well as Brian Scholl and two anonymous reviewers for their very helpful insights.

References

Adler, S. A., Gerhardstein, P., & Rovee-Collier, C. (1998). Levels-of-processing effects in infant memory? *Child Development*, *69*(2), 280–294.

Bhatt, R. S., & Rovee-Collier, C. (1997a). Dissociation between features and feature relations in infant memory: effects of memory load. *Journal of Experimental Child Psychology*, *67*, 69–89.

Bhatt, R. S., & Rovee-Collier, C. (1997b). Perception and 24-hour retention of feature relations in infancy. *Developmental Psychology*, *30*(2), 142–150.

Carey, S., & Xu, F. (2001). Infants' knowledge of objects: beyond object files and object tracking. *Cognition*, *80*, 179–213.

Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*, 87–185.

Dehaene, S., & Changeux, J. P. (1993). Development of elementary numerical abilities: a neuronal model. *Journal of Cognitive Neuroscience*, *5*, 390–407.

Ericsson, K. A., Chase, W. G., & Faloon, S. (1980). Acquisition of a memory skill. *Science*, *208*, 1181–1182.

Feigenson, L (2002). *The representations underlying more/less comparisons*. Paper presented at the International Conference on Infant Studies, Toronto, Canada.

Feigenson, L., & Carey, S. (2003). Tracking individuals via object-files: evidence from infants' manual search. *Developmental Science*, *6*, 568–584.

Feigenson, L., Carey, S., & Hauser, M. (2002). The representations underlying infants' choice of more: object files versus analog magnitudes. *Psychological Science*, *13*(2), 150–156.

Feron, J., Streri, A., & Gentaz, E (2002). *Numerical intermodal transfer from touch to vision by 5-month-old infants*. Poster presented at the International Conference on Infant Studies, Toronto, Canada.

Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition*, *44*, 43–74.

Gallistel, C. R., & Gelman, R. (2000). Non-verbal numerical cognition: from reals to integers. *Trends in Cognitive Science*, *4*(2), 59–65.

Gerhardstein, P., & Rovee-Collier, C. (2002). The development of visual search in infants and very young children. *Journal of Experimental Child Psychology*, *81*, 194–215.

Gulya, M., Rovee-Collier, C., Galluccio, L., & Wilk, A. (1998). Memory processing of a serial list by young infants. *Psychological Science*, *9*(4), 303–307.

Gulya, M., Sweeney, B., & Rovee-Collier, C. (1999). Infants' memory processing of a serial list: list length effects. *Journal of Experimental Child Psychology*, *73*, 72–91.

Halberda, J., Simons, D., & Wetherhold, J (2003). Change-detection reveals the top-down impenetrability of visual short-term memory, and the 3-item limit of parallel attention. Manuscript submitted for publication.

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

Kobayashi, T., Hiraki, K., & Hasegawa, T (2002). *Intermodal numerical correspondences in 6-month-old infants*. Poster presented at the International Conference on Infant Studies, Toronto, Canada.

Lipton, J., & Spelke, E. (2003). Origins of number sense: large number discrimination in six-month-old infants. *Psychological Science*, *14*, 396–401.

Miller, G. M. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, *63*, 81–97.

Palmer, S. E., & Rock, I. (1994). Rethinking perceptual organization: the role of uniform connectedness. *Psychonomic Bulletin and Review*, *1*, 29–55.

Rensink, R. A. (2000). Visual search for change: a probe into the nature of attentional processing. *Visual Cognition*, *7*, 345–376.

Rovee-Collier, C. (1999). The development of infant memory. *Current Directions in Psychological Science*, *8*(3), 80–85.

Rovee-Collier, C., Bhatt, R. S., & Chazin, S. (1996). Set size, novelty, and visual pop-out in infancy. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(5), 1178–1187.

Rovee-Collier, C., Hankins, E., & Bhatt, R. S. (1992). Textons, visual pop-out effects, and object recognition in infancy. *Journal of Experimental Psychology: General*, *121*(4), 435–445.

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

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 multiple object tracking. *Cognition*, *80*, 159–177.

Scholl, B. J., & Xu, Y. (2001). The magical number 4 in vision. *Behavioral and Brain Sciences*, *24*, 145–146.

Simon, H. A., & Chase, W. G. (1973). Skill in chess. *American Scientist*, *61*, 393–403.

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

Spelke, E. S. (1994). Initial knowledge: six suggestions. *Cognition*, *50*, 431–445.

Starkey, P., Spelke, E., & Gelman, R. (1990). Numerical abstraction by human infants. *Cognition*, *36*, 97–127.

Strauss, M. S., & Curtis, L. E. (1981). Infant perception of numerosity. *Child Development*, *52*, 1146–1152.

Trick, L., & Pylyshyn, Z. W. (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.

Wertheimer, M (1923). Untersuchungen zur Lehre von der Gestalt, II. *Psychologische Forschung, 4,* 301–350. Condensed translation published as Laws of organization in perceptual forms, in Ellis, W. D (1938). *A sourcebook of gestalt psychology* (pp. 71–88). New York: Harcourt, Brace.

Whalen, J., Gallistel, C. R., & Gelman, R. (1990). Nonverbal counting in humans: the psychophysics of number representation. *Psychological Science*, *10*, 130–137.

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

Wynn, K. (1998). Psychological foundations of number: numerical competence in human infants. *Trends in Cognitive Sciences*, *2*, 296–303.

Wynn, K., Bloom, P., & Chiang, W-C. (2002). Enumeration of collective entities by 5-month old infants. *Cognition*, *83*(3), B55–B62.

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

Xu, F., Spelke, E. S., & Goddard, S (2003). Number sense in human infants. Manuscript submitted for publication.

Yantis, S., & Johnson, D. (1990). Mechanisms of attentional priority. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 812–825.