Abtract—Do genuinely numerical computational abilities exist in infancy? It has recently been argued that previous studies putatively illustrating infants’ ability to add and subtract tapped into specialized object-tracking processes that apply only with small numbers. This argument contrasts with the original interpretation that successful performance was achieved via a numerical system for estimating and calculating magnitudes. Here, we report that when continuous variables (such as area and contour length) are controlled, 9-month-old infants successfully add and subtract over numbers of items that exceed object-tracking limits. These results support the theory that infants possess a magnitude-based estimation system for representing numerosities that also supports procedures for numerical computation.

A growing number of studies suggests that humans possess a system for nonverbal representation of numerical magnitudes. Very young infants (and in some cases, newborn infants) discriminate small numbers of syllables and tones (Bijeljac-Babic, Bertoncini, & Mehler, 1993; vanMarle & Wynn, 2003), moving objects and collections of objects (van Loosbroek & Smitsman, 1990; Wynn, Bloom, & Chiang, 2002), and simple dots (Antell & Keating, 1983; Starkey & Cooper, 1980). Furthermore, infants as young as 5 months of age have been shown to conduct arithmetical operations such as addition and subtraction over these represented quantities. Wynn (1992) showed 5-month-olds a Mickey Mouse doll, covered this toy with an occluder, and placed another doll behind the occluder (creating a 1 + 1 scenario). A second group of infants viewed an analogous subtraction operation, seeing one doll removed from an occluded pair of Mickey dolls (2 − 1). In both scenarios, the occluder then dropped to reveal either one object or two objects. Infants tended to look longer to the incorrect outcome than the correct outcome; that is, the addition group looked longer at the outcome of one doll than two dolls, whereas the reverse pattern held for the subtraction group. A follow-up experiment suggested that infants did not simply expect “more” or “less” than the initial number seen, but rather expected exactly the correct number of dolls.

Wynn (1992, 1995) concluded that very young infants have a “number sense” similar to that found in other animals. This primitive capacity allows organisms to represent and manipulate numerical values via a magnitude-estimation system. Recent studies on this capacity in adult humans (e.g., Whalen, Gallistel, & Gelman, 1999) support the existence of a magnitude-based estimation system (such as the accumulator mechanism proposed by Meck & Church, 1983). Several researchers have proposed that this same mechanism underlies infants’ numerical abilities (Gallistel & Gelman, 1992; Wynn, 1995).

Recently, however, it has been suggested (Simon, 1997; Uller, Carey, Huntley-Fenner, & Klatt, 1999) that infants’ numerical competence is a result not of a numerical representation system, but rather of an automatic system for tracking and reasoning about individual objects in the world, via “object files” or other object-tracking mechanisms (e.g., Kahneman, Treisman, & Gibbs, 1992; Scholl & Leslie, 1999; Trick & Pylyshyn, 1994). In this account, each individual object in an array is represented by an opened “file.” Numerical differences between two sets in this system are detected by a one-to-one correspondence function. By this theory, infants’ longer looking times to incorrect outcomes in addition and subtraction paradigms are due to a mismatch between the object files stored in memory and the objects present visually; no one-to-one correspondence is possible. Unlike a magnitude-representation system, whose job is to represent numerosity, object files are largely used to track objects’ locations and paths of movement.
stringent capacity limits for object files; most research shows a limit of three to four objects to which object files may be assigned at any one time (e.g., Pylyshyn & Storm, 1988; for a review, see Trick & Pylyshyn, 1994).

Feigenson, Carey, and Hauser (2002) directly tested whether the numerical competence found in number-discrimination and addition and subtraction tasks is actually the result of an object-based system. They presented 10- and 12-month-old infants with two amounts of graham crackers, placed one at a time into separate containers, and noted which container the children preferred to crawl to. The infants correctly chose the container with more crackers when the number of crackers in each container was less than four (e.g., one vs. two or two vs. three). When one container held four or more crackers (such as three vs. six, or even one vs. four), the infants chose randomly. Thus, only when the number of crackers involved was within the limits of an object-file system were the children able to reach a judgment of “greater than.” The authors concluded from this failure with large numbers that the infants were not relying on an analog-magnitude representational system. Instead, this set-size signature indicates that the infants used a system dedicated to tracking small numbers of objects (Feigenson, Carey, & Hauser, 2002; see also Hauser, Carey, & Hauser, 2000, for similar findings with adult monkeys). This raises the possibility that previous studies of addition and subtraction abilities (e.g., Wynn, 1992) tapped object-tracking processes rather than numerical cognition, leaving open the question of whether genuinely numerical computational abilities are present in infancy.

One way to address this question is to study infants’ performance in situations that exceed the limits of object-tracking systems. To date, very few studies have investigated infants’ ability to perform numerical operations over large numbers. The studies that have been done have focused on the operation (defined as such by Gallistel & Gelman, 1992) of ordering numerosities according to relative amount. Evidence for this ordinal competence is mixed; the work already mentioned (Feigenson, Carey, & Hauser, 2002) found that infants were unsuccessful with greater-than/smaller-than comparisons involving large numbers, whereas Brannon (2002) found that 11-month-old infants (but not 9-month-olds) appear to successfully order values. Some evidence for representing and computing over large numbers can be found in recent research with animals. Brannon, Wusthoff, Gallistel, and Gibbon (2001) found that pigeons are able to discern correct from incorrect outcomes of a subtraction task. Flombaum, Junge, and Hauser (2003), using a looking-time measure, found that rhesus macaques were capable of inferring the outcomes of addition and subtraction of large numbers of objects (4 + 4 = 8 or 4). There is also evidence for nonverbal magnitude estimation and manipulation from work with adult humans (Barth, Kanwisher, & Spelke, 2003; Whalen et al., 1999).

Other recent work has begun to examine large-number representations during infancy. Xu and Spelke (2000) habituated 6-month-old infants to displays containing either 8 dots or 16 dots. At test, the infants were given alternating trials of 8 dots and 16 dots. The infants disambiguated to the novel number, showing successful discrimination of these larger values. In a second experiment using an identical design, infants failed to discriminate 8 from 12. Subsequent research (Xu, Spelke, & Goddard, 2003) illustrated that infants this same age also succeeded at discriminating 16 from 32, but failed to discriminate 16 from 24. These results indicate that young infants can represent number independently of tracking individual objects, and that their number representations are highly inaccurate. Interestingly, the discrimination function obtained in these studies displays the Weber-fraction signature of scalar variance, in which it is the ratio of two values, rather than their absolute difference, that determines their discriminability. This scalar variability is a hallmark of the accumulator model. Chiang and Wynn (2000) also looked at infants’ representations of large numbers of objects, finding that infants could successfully reason about the “magical disappearance” of a large number (specifically, a collection) of objects provided the objects’ contours and individual identities were made initially salient.

The purpose of the current study was to establish whether infants possess procedures for numerical computation, using the domain of large numbers as a test case. Because infants’ performance in addition and subtraction studies to date may have reflected the operation of object-tracking processes rather than numerical processing, there is no compelling evidence that infants can successfully operate over numerical magnitudes. If infants’ magnitude-estimation system (believed to underlie discrimination of large numbers) supports numerical manipulations and operations, then infants should be successful with large-number operations. Accordingly, we presented infants with computerized displays of the large-number addition and subtraction operations 5 + 5 and 10 − 5.

METHOD

Participants
The final sample consisted of twenty-six 9-month-old infants with a mean age of 9 months 4 days (range: 8 months 14 days to 9 months 25 days). Participants were recruited from the Southeastern Connecticut community by way of commercially purchased lists and hospital sign-up sheets. An additional 16 infants were tested but excluded because they completed fewer than two full test pairs (n = 7), were fussy (n = 3), failed to look at the monitor (n = 5), or received insufficient occluder experience (n = 1; see Scoring). Data from 1 outlier (> 2 SDs from mean difference score) were discarded and replaced with an additional participant’s data.

Design and Procedure
Participants were divided into two groups—addition and subtraction. Each group had approximately equal numbers of males
and females (7 males and 6 females in the addition group, 6 males and 7 females in the subtraction group), and within each group, the correct and incorrect outcomes were presented first approximately equally often (12 incorrect first, 14 correct first).

Infants sat in a car seat positioned 75 cm from a computer monitor, surrounded by black curtains that concealed the experimenters and equipment. Each infant's parent sat next to him or her, but faced away from the screen so as not to influence the child's looking behavior. An observer watching through a hole in the curtain surreptitiously coded the infants' looking time online using a computerized key-press program. The infants first saw occluder-familiarization and outcome-familiarization movies (see Stimuli). Each infant was then shown the test movies, alternating between correct and incorrect outcomes, with each alternating set forming a test pair. A trial ended when the child looked away from the monitor continuously for 2 s, and a curtain between the infant and computer was lowered while the next trial was queued on screen. At the beginning of each trial, the curtain was raised, and the experimenter said, “Watch the movie!” while squeaking a toy. During the test movies, the toy was again squeaked right before the critical subtraction or addition event (as the screen was raised to cover the first set of objects) to ensure that the infants paid attention. A trial was included in the analyses only if the infant watched the initial occlusion of a set of objects and then the addition or subtraction of a second set of objects (approximately 5 s of video).

Stimuli
Events were presented on a computer monitor. All animations were created using Infini-D animating software and rendered into Quicktime movie clips. All movies were rendered against a black background, with a 2-cm purple computerized occluder at the top of the screen. The infants were shown three types of movies: the occluder-familiarization movie, which was designed to give them experience with the nature of the objects in the study; outcome movies, which were designed to familiarize them to the sight of the outcomes; and the test events themselves. (The movies can be viewed on the Web at http://pantheon.yale.edu/~kcm32/research/multimedia.html.)

Previous research (Clearfield & Mix, 1999; Feigenson, Carey, & Spelke, 2002) suggests infants may be sensitive to continu-ous-extent variables that covary with number (such as area or contour length) and may utilize these variables in generating expectations about the outcome of an event. Consequently, we equated the test outcomes of 5 and 10 objects for continuous-extent values (each outcome had a summed area of 39.45 cm$^2$ and contour length of 90 cm). The objects presented varied in size. The average area of a single object in the display showing 5 as the outcome was 7.89 cm$^2$ (contour length of 18 cm), and the average area of a single object in the display showing 10 as the outcome was 3.95 cm$^2$ (contour length of 9 cm). The objects that underwent the operations shrank and grew constantly and individually until they moved over to the right side of the screen (where they were covered). By the time the occluder was raised, the objects were a uniform size. Thus, the objects that comprised the outcomes were very different (e.g., variable in terms of their rectangular dimensions) from the objects that were occluded.

Occluder-Familiarization Movie
In this movie, a rectangle (4.93 cm × 1.07 cm) moved onto the left side of the screen from behind the purple occluder. The rectangle rotated fully three times in place and moved to the right side of the screen (all the while shrinking and growing). The mean area of the rectangle across its changes was approximately 5.3 cm$^2$, with minimum dimensions of 2.24 × 1.06 cm and maximum dimensions of 11.04 × 0.96 cm. Once the rectangle was on the right side of the screen, a white computerized occluder appeared from the bottom of the monitor and covered it. After 250 ms, the occluder moved off screen to reveal the object. The object was shrinking, growing, and rotating slightly as it became uncovered. This occlusion event occurred five times.

Outcome-Familiarization Movies
These two movies (one animation per numerical outcome) allowed us to examine pretest preferences for one type of outcome over the other, and also served to familiarize the infants to the perceptual features of the outcomes. The movies showed a white computerized occluder on the right side of the screen. After the experimenter got the child's attention by squeaking a toy, the occluder moved off screen to reveal 5 or 10 rectangles. These rectangles were the exact same displays as the testing outcomes. All infants saw both outcomes.

Test Movies
For infants in the addition group, the test movies (see Fig. 1) showed a set of 5 rectangles that stretched and fell down from behind the purple occluder, and moved over to the right side of the screen. There they were covered by a white computerized occluder. Five other objects emerged serially from behind the purple occluder and, after lining up on the left side of the screen, moved under the white occluder, which then moved off screen to reveal an incorrect outcome of 5 items or a correct outcome of 10. For infants in the subtraction group, the test movies (see Fig. 1) showed 10 objects that stretched down onto the left side of the screen from behind the purple occluder. After they moved to the right side of the monitor, the white occluder covered them up. Five objects then emerged from behind the white occluder and moved serially off the screen. The white occluder then dropped to reveal either an incorrect outcome of 10 or a correct outcome of 5.

Scoring
Infants who did not see at least one example of an occlusion during the familiarization movies were excluded from the
experimental analyses. The criterion for the minimum look to
the outcomes was 2 s; trials in which this criterion was not met
were repeated, as were any trials in which the infant did not look
at the critical addition or subtraction event.

Looking time was measured by an observer who was hidden
behind a curtain and was unaware of the infant’s assigned op-
eration group (addition or subtraction). A second experimenter,
also naive to the subject’s operation group, reviewed video
footage of a subset of participants (20%) and measured the
looking times. These times were found to be very highly cor-
related to the on-line timing ($r = .99$), and thus all data analyses
were performed using results from the primary coder.

RESULTS

Baseline looking times to outcomes of 5 or 10 objects (9.12 and
9.68 s, respectively) were not significantly different, $t(1,
25) = -0.48, p = .636$. Preliminary analyses revealed no ef-
eects of sex (male, female), trial order (correct first, incorrect
first), or test pair (first, second, third); our subsequent analyses
collapsed over these variables. A 2 (outcome: 5 objects vs. 10
objects) × 2 (operation group: addition vs. subtraction) re-
peated measures analysis of variance on infants’ mean looking
times at test (see Fig. 2) demonstrated a significant interaction,
$F(1, 25) = 6.693, p = .012, \eta^2 = .22$. Infants who saw an add-
dition operation looked longer at the outcome of 5 (10.28 s) than
at the outcome of 10 (7.35 s; mean looking-time differ-
ence $= 2.93$ s), whereas infants who saw a subtraction operation
looked longer at the outcome of 10 (9.13 s) than at the outcome
of 5 (8.00 s; mean looking-time difference $= 1.13$ s).

For nonparametric analyses, infants were categorized as
preferring the incorrect test outcomes if their looking-time
difference (average of incorrect trials – average of correct trials)
was positive; a negative difference score characterized a pref-
erence for correct movies. Twenty of 26 infants (10 in each
group) preferred the incorrect movie, and a $2 \times 2$ chi-squared
contingency test (Habituation Group × Test Trial Preference)
was significant, $\chi^2(1, N = 26) = 7.54, p < .01$. A significant

---

**Addition**

- five objects drop down
- the occluder rises
to cover them
- five additional objects emerge
  and go behind the occluder

**Subtraction**

- ten objects drop down
- the occluder rises
to cover them
- five objects move from
  behind the occluder, and go offscreen

-the occluder drops to reveal an outcome of either:
  - five objects
  (for half of the trials)
  - ten objects
  (for the other half of the trials)

**Fig. 1.** Schematic of the addition and subtraction test movies.

---

**Fig. 2.** Looking time to the outcome as a function of test pair (first, sec-
ond, or third), operation group (addition, subtraction), and number of
outcome objects (5, 10). Infants in the addition group saw a $5 + 5$ sce-
enario, and infants in the subtraction group saw a $10 − 5$ scenario.
effect was also found in each operation group independently, as 10 of 13 infants in both the addition and subtraction groups preferred the incorrect movie; one-tailed sign tests yielded a p value of .049 for each group, z(12) = 1.66.

**DISCUSSION**

In this study, infants were given correct and incorrect outcomes to a mathematical operation (5 + 5 or 10 − 5). Infants who saw the addition operation looked longer to an outcome of 5 than to an outcome of 10, and infants who saw the subtraction operation looked longer to an outcome of 10 than to an outcome of 5. This pattern was not driven by baseline preferences for one of the two outcome displays, or by operations over continuous-extent variables. We conclude that the infants were performing numerical computations over represented magnitudes.

This is the first evidence of preverbal children adding and subtracting over quantities exceeding the limits of the object-file system. The fact that infants can compute outcomes of addition and subtraction over large numbers indicates they have procedures for numerical operations independent of object-tracking processes. Specifically, these results suggest that infants can numerically operate over magnitude representations of large numbers. Humans possess an early system that supports numerical combination and manipulation, even before the powerful tool of language develops.

Cohen and Marks (2002) recently posited that the results found in addition and subtraction tasks can be readily explained by a dual-process model in which infants are expressing solely perceptual preferences. Central to this model are two common looking-time patterns, in which infants prefer (a) familiar sets and (b) larger amounts. It is well documented that complex, briefly presented, or superficially processed stimuli will drive looking time to a familiar stimulus over a novel one, and this heightened processing of the familiar stimulus is especially apparent in young infants (such as the 5-month-old group whose data elicited this model). In addition, the more objects present on the screen, the more there is to process, and therefore infants tend to look longer at larger than at smaller sets. Together, these two processes would predict the same pattern as found in traditional preverbal addition and subtraction studies (e.g., Wynn, 1992).

It should be noted that attention, memory, and looking time surely interact in any study using a looking-time paradigm. However, one of the central findings to come out of the visual recognition memory literature is that the degree to which a novel stimulus is preferred over a familiar stimulus (a preference thought to reflect speed of processing) dramatically increases across the first year of life (e.g., Colombo, Mitchell, & Horowitz, 1988; Rose, 1983; Rose, Feldman, & Jankowski, 2001). Therefore, one cannot make the claim that the infants in the present study (who were almost twice as old as 5-month-olds) showed the purported familiarity preference that led Cohen and Marks (2002) to include this tendency in their dual-process model. Additionally, it is unclear what is “familiar” in the present context. The sets to be manipulated changed size and rectangular dimensions because we wanted to induce a percept of variability. The sole purpose of this variability was to prevent the formation of a representation of a particular perceptual amount. Finally, the general tendency of infants to look more at one particular amount than at another was not a concern in the present study; we found no significant baseline differences in looking times to the two outcomes. Given these familiarity controls and baseline looking-time patterns, the present findings must reflect more than a simple layering of perceptual responses.

These data do not speak directly to the issue of what systems are employed in small-number addition and subtraction experiments. However, they do establish a key piece of evidence that the proposed nonverbal magnitude-based estimation system is deeply continuous in nature. It is present in various animal species, is present very early in human infancy, and extends all the way through adulthood (Barth et al., 2003; Brannon et al., 2001; Xu & Spelke, 2000). These data lend credence to the idea that “humans innately possess the capacity to perform simple arithmetical calculations, which may provide the foundations for the development of further arithmetical knowledge” (Wynn, 1992, p. 750).

As noted in the introduction, the perception of small numbers of visual things inevitably recruits object-based processes. Therefore, disentangling the contributions of object-tracking and magnitude-representation mechanisms to performance in small-number tasks is a challenging enterprise. It is our view that an “either-or” dichotomy regarding this set of cognitive systems would be mistaken. Instead, we suggest an account similar to that espoused by Xu (2003): There exist both an object-tracking system that represents small numbers of visually presented objects and a number system that can represent both large and small approximate magnitudes of items presented in any modality. When an observer is faced with a small array of objects, both systems are likely engaged. In modalities where there are not “objects” as such, one can examine representations of numerosity without the confound of objecthood. For example, auditory experiments have found both small-number (Bijeljac-Babic et al., 1993; vanMarle & Wynn, 2003) and large-number (Lipton & Spelke, 2003) discrimination capabilities. These studies indicate that object-tracking processes are insufficient to account for the entirety of infants’ number-discrimination abilities. Although some conditions may favor the utilization of object-based cognitive structures over number-based ones, the number sense—the ability to nonverbally represent approximate numerosities and do computations over these representations—is present even in infancy. The current findings are a testament to this independent numerical prowess.

The theory that large-number representations and manipulations are supported by a magnitude-estimation system suggests
follow-up studies to explore the limits to infants’ computing ability. As noted in the introduction, the Weber-fraction signature of scalar variance is a telltale sign of magnitude-estimation systems. The 9-month-olds in the current study could competently discriminate the outcomes of 5 and 10 rectangles, outcomes that differ by a 2:1 ratio. If infants’ number representations are supported by a magnitude-estimation system, infants should fail our task if the two outcome values differ by a much smaller ratio (e.g., 10 vs. 8). There are also questions of a developmental nature to address. The discrimination function appears to improve quickly through infancy. Lipton and Spelke (2003) found that 9-month-olds, but not 6-month-olds, could discriminate values differing by a 2:3 ratio (e.g., 8 vs. 12). Given this increasing level of representational precision with age, researchers should also find corresponding age differences in computational precision. By establishing the extents and limits of this system, investigators can map out a vital cognitive capacity.

Acknowledgments—The authors are grateful to Paul Bloom, Erik Cheries, Jonathan Flombaum, Valerie Kuhlmeier, and Kristy vanMarle for their comments on earlier drafts. We would like to thank the parents who brought in their children to participate and our laboratory coordinator, Cindy Decoste, without whom this research would not be possible. This work was supported by a grant to the second author from the National Science Foundation.

REFERENCES


(Received 9/15/03; Revision accepted 11/3/03)