Developmental Science 8:1 (2005), pp 88–101

PAPER

Number sense in human infants

Fei Xu,[1, 3] Elizabeth S. Spelke[2] and Sydney Goddard[1]

1. *Northeastern University, Boston, USA*
2. *Harvard University, Cambridge, USA*
3. *Now at the University of British Columbia, Vancouver, Canada*

Abstract

Four experiments used a preferential looking method to investigate 6-month-old infants' capacity to represent numerosity in visual-spatial displays. Building on previous findings that such infants discriminate between arrays of eight versus 16 discs, but not eight versus 12 discs (Xu & Spelke, 2000), Experiments 1 and 2 investigated whether infants' numerosity discrimination depends on the ratio of the two set sizes with even larger numerosities. Infants successfully discriminated between arrays of 16 versus 32 discs, but not 16 versus 24 discs, providing evidence that their discrimination shows the set-size ratio signature of numerosity discrimination in human adults, children and many non-human animals. Experiments 3 and 4 addressed a controversy concerning infants' ability to discriminate large numerosities (observed under conditions that control for total filled area, array size and density, item size and correlated properties such as brightness: Brannon, 2002; Xu, 2003b; Xu & Spelke, 2000) versus small numerosities (not observed under conditions that control for total contour length: Clearfield & Mix, 1999). To investigate the sources of these differing findings, Experiment 3 tested infants' large-number discrimination with controls for contour length, and Experiment 4 tested small-number discrimination with controls for total filled area. Infants successfully discriminated the large-number displays but showed no evidence of discriminating the small-number displays. These findings provide evidence that infants have robust abilities to represent large numerosities. In contrast, infants may fail to represent small numerosities in visual-spatial arrays with continuous quantity controls, consistent with the thesis that separate systems serve to represent large versus small numerosities.

A wealth of evidence indicates that humans and other animals have a sense of approximate numerical magnitudes, and that this sense depends on a cognitive system that emerges in humans by five years of age (Dehaene, 1997; Gallistel, 1990, Gallistel & Gelman, 2000; Huntley-Fenner, 2001; Temple & Posner, 1998). The signature property of the system underlying number sense in human adults, children and non-human animals is *scalar variability* (e.g. Cordes, Gelman, Gallistel & Whalen, 2001; Van Oeffelen & Vos, 1982; Mechner, 1958): the error in numerosity representations is proportional to numerical magnitude, and therefore discriminability between two numerosities depends on their ratio. To date, however, few studies have investigated human infants' representations of large numerosities, and none has investigated whether infants' representations of number show the set-size ratio signature of the mechanisms of number sense in other populations.

The question whether infants' number representations show the same signature properties as the number representations of children and adults is highly important in light of current debates over the developmental origins of number sense. Although infants have long been known to discriminate between small numbers of objects (e.g. Antell & Keating, 1983; Starkey & Cooper, 1980; Starkey, Spelke & Gelman, 1983; Strauss & Curtis, 1981; Treiber & Wilcox, 1984), the bases of this discrimination continue to be debated (Mix, Huttenlocher & Levine, 2002). In early studies, infants were found to discriminate displays of two versus three visual elements when the elements either had the same size (e.g. Starkey & Cooper, 1980) or varied in size randomly (e.g. Treiber & Wilcox, 1984). In such experiments, however, changes in number are correlated with changes in a variety of continuous quantities including total filled area, brightness, contour length in displays of two-dimensional forms, and the total filled volume in displays of three-dimensional objects. More recent studies, testing infants' small-number discrimination in two-dimensional displays with strict controls for total contour length or in three dimensional

Address for correspondence: Fei Xu, Department of Psychology, 2406-2136 West Mall, Kenny Building, University of British Columbia, Vancouver, B.C. V6T 1Z4, Canada; e-mail: fei@psych.ubc.ca

displays with strict controls for total volume, have found that infants respond to the latter variables and not to number (Clearfield & Mix, 1999; Feigenson, Carey & Spelke, 2002).

Some experiments have circumvented the problem of correlated spatial variables such as contour length and volume by investigating infants' discrimination of small numbers of visible events (jumps of a puppet), auditory events (syllables in a spoken word), or visible collections (groups of forms undergoing distinctive rigid motions) (Bijelbac-Babic, Bertoncini & Mehler, 1991; Sharon & Wynn, 1998; Wynn, Bloom & Chiang, 2002). Infants have been found to discriminate between small numbers of actions, syllables and collections, despite the absence of any correlation of number with contour length or volume. In these studies, however, number was correlated with other continuous variables such as the total amount of motion in the event sequences (the puppet covered a greater total distance when it performed more jumps), the total amount of sound in the words (greater for words with more syllables), and the variability of motion in the display (the larger number of moving groups presented a greater array of relative motions). After more than two decades of study, therefore, there is still no consensus over whether infants discriminate between small numbers of elements on the basis of numerosity.

Even if infants were found to represent small numerosities when all correlated continuous variables were controlled, the basis of this ability would not be clear. Some investigators have proposed that the capacity to discriminate small numbers of elements depends on the same sense of approximate number that human adults and non-human animals use to discriminate larger numerosities: discrimination of small numbers only appears to be exact because changes from *one* to *two* to *three* fall above the critical ratio-limit on number discrimination in most tasks (Cordes *et al.*, 2001; Dehaene, 1997; Gallistel & Gelman, 2000; Wynn, 1998). Others have proposed that small-number discrimination depends on a special process of 'subitizing' (e.g. Balakrishnan & Ashby, 1992; Mandler & Shebo, 1982; Starkey & Cooper, 1980), linked to mechanisms for representing individual objects (Carey, 2001; Scholl, 2001; Simon, 1997; Trick & Pylyshyn, 1994). Consistent with the latter claim, Starkey and Cooper (1980) reported that 6-month-old infants' number discrimination fails to show the set-size ratio signature found in studies of number sense in older children and adults: infants successfully discriminated two from three discs but failed to discriminate four from six discs: sets that differed by the same ratio but lay outside the subitizing range (see also Feigenson, Carey & Hauser, 2002). Because these studies failed to control for continuous quantities, however, their relevance to infants' number discrimination can be questioned.

To investigate whether human infants have the same sense of approximate numerical magnitudes that older children and adults do, therefore, it is necessary to investigate infants' discrimination of larger numerosities under conditions that control for all the continuous quantities to which infants might be sensitive. If number sense originates in infancy, then infants should discriminate large as well as small numerosities when presented with arrays in which all continuous variables are controlled. Moreover, infants' numerosity discrimination should show the set-size ratio signature found in adults, children and non-human animals.

A number of recent studies have begun to test these predictions (Brannon, 2002; Brannon, Abbott & Lutz, in press; Lipton & Spelke, 2003; Xu, 2003b; Xu & Spelke, 2000). Using a preferential looking method relying on infants' tendency to look longer at novel arrays, Xu and Spelke (2000; see also Brannon, *et al.*, in press) found that 6-month-old infants discriminated arrays of eight discs from arrays of 16 discs under conditions that controlled for the continuous variables of display size, element density, element size, total filled area, and correlated properties such as surface brightness and texture (a detailed description of these controls is given below). Xu (2003b) found that 6-month-old infants also discriminated arrays of four discs from eight discs when total filled area or contour length was controlled for. Using a head-turn preference method, Lipton and Spelke (2003) found that 6-month-old infants discriminated between sequences of eight versus 16 sounds when the continuous variables of element length, sequence length, sequence rate and total amount of sound were controlled. In further research, in contrast, 6-month-old infants failed to discriminate between visual arrays presenting eight versus 12 discs (Xu & Spelke, 2000) or between auditory sequences presenting eight versus 12 sounds (Lipton & Spelke, 2003). Lipton and Spelke (in press) also established that infants' sensitivity depends on the ratio and not the absolute difference between two numbers, since infants succeeded in discriminating four versus eight sounds but failed at four versus six sounds. These findings provide evidence that infants can indeed discriminate between large sets of elements on the basis of numerosity and that their numerical discriminations are highly imprecise (ratio limit between 2:3 and 1:2), consistent with the hypothesis that a sense of approximate numerical magnitudes emerges in infancy.

The findings, however, leave several questions unanswered. Although previous studies have established that successful discrimination of numerosities in infants depends on the ratio of the set sizes, so far all studies have used stimuli with a narrow range: 4–16 discs or sounds (Lipton & Spelke, 2003, in press; Xu, 2003; Xu & Spelke, 2000). Studies with other animals suggest that they can discriminate numerosities larger than 16 (Gallistel, 1990). Here we further explore the ratio limit in infants with numerosities twice as high as any previously tested, i.e. 16 versus 32. It is possible that infants' numerical discrimination shows a different pattern of variation across the number range: infants' large-number representations, like their small-number representations, may even show a set size limit and break down altogether for numerosities above 16. Experiments 1 and 2 investigated 6-month-old infants' discrimination between arrays of 16 versus 32 discs and arrays of 16 versus 24 discs.

The second unanswered question raised by past studies concerns the apparent discrepancy between the findings of studies of large-number and small-number discrimination. We have noted that infants have shown successful discrimination among sets of one, two, or three elements only when numerosity was confounded with other, continuous variables such as the total contour length, total volume of material, or the amount or variability of motion in the display. When total contour length or volume were controlled, infants failed to discriminate between displays of one, two, or three elements (Clearfield & Mix, 1999; Feigenson, Carey & Spelke, 2002). What explains the contrast between these results and those that presented larger numerosities and controls for the continuous quantities of total filled area, total contour length, display size and density, and element size (Brannon, 2002; Xu, 2003b; Xu & Spelke, 2000)? It is possible that infants are sensitive to numerosity only when the sets they encounter are large; when presented with three or fewer objects, infants may represent the objects as individuals, not as a set with a cardinal value (Feigenson, Carey & Spelke, 2002; Simon, 1997; Spelke, 2000). Experiments 3 and 4 were conducted using numerosities and continuous quantity controls complementary to those of past research. Experiment 3 tested infants' large number discrimination with controls for contour length, measured as in Clearfield and Mix's (1999) studies of small-number discrimination.[1] Experiment 4 tested infants' small-number discrimination with controls for total filled area, display size and density, and element size, as in past studies of large-number discrimination.

Experiment 1

Method

Six-month-old infants were presented with a succession of visual arrays containing either 16 or 32 discs. For a given infant, all the arrays contained the same number of discs, randomly spread over the same total display area; the sizes and positions of the discs varied from one array to the next. As in Xu and Spelke's (2000) experiments, the arrays with the two different numerosities (16 and 32) were equated in overall size and in total filled area, and so they were equal in brightness. Because these variables were equated, the arrays with the larger numerosity presented discs that were half as large, on average, and twice as dense as those with the smaller numerosity (Figure 1). The 16- or 32-element arrays were presented to each infant until his or her looking time declined to a criterion of habituation, and then all the infants were presented with six new test arrays containing 16 and 32 discs in alternation. Arrays of the two numerosities were now equated in element size and density (and therefore appeared to present the same visible texture) and varied in total array size and filled area.

These stimulus variations allowed us effectively to distinguish responses to numerosity from responses to a host of continuous quantitative variables, including display size, element size and density, total filled area, and the correlated properties of surface brightness and texture. If infants responded to any of these variables then the infants in the two habituation groups should have shown no preference between the two test displays, because every continuous variable that distinguished the two sets of habituation displays was equated across the two sets of test displays, and vice versa. In contrast, if infants were sensitive to the numerical differences between the arrays, then the infants in each habituation group should have looked longer at the test displays with the novel numerosity.

Participants

Participants were 16 full-term infants (eight males, eight females) ranging in age from 5 months, 15 days to 6 months, 15 days (mean age = 6 months, 0 days). An additional six infants were eliminated from the sample because of fussiness or parental interference. All infants were recruited by obtaining their birth records from town halls in the greater Boston area and subsequently contacting their parents by mail and telephone.

Apparatus and stimuli

The apparatus was the same as in Xu and Spelke (2000). Infants sat in an infant seat facing a well-illuminated

---

[1] Because volume is not defined for two-dimensional displays, it is not possible to test for the effects of variations in volume on infants' large-number discrimination using displays like those of the present experiments.

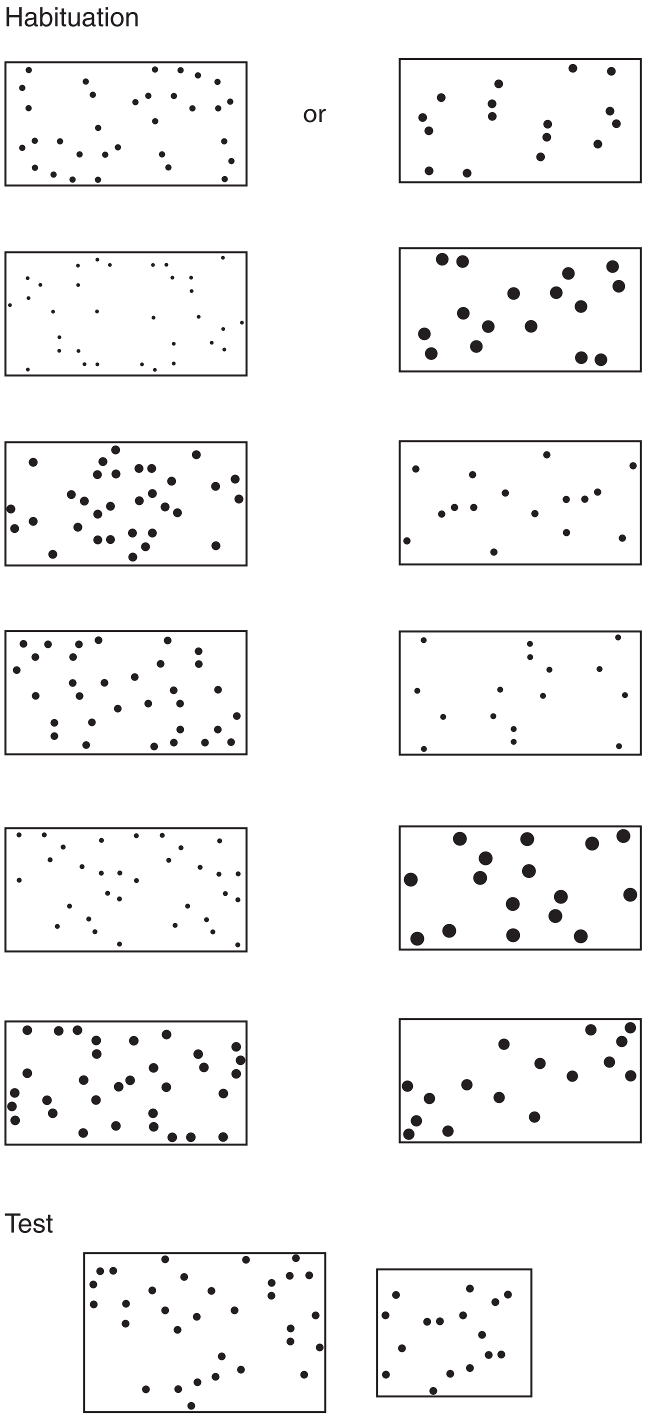

Figure 1 *Selected habituation and test displays for Experiment 1.*

puppet stage surrounded by black curtains. A camera focused on the infant through a small hole below the stage center; it was connected to a television monitor and a VCR in a corner of the room, on which an observer watched the infant and recorded his or her looking times by a button box connected to a computer. The observer was blind both to the infant's condition and to the displays viewed on any given trial. All infants were observed offline by a second observer who also was blind to these variables. Interscorer reliability averaged 92%. A display camera was placed behind the infant to record the sequence of the displays. A parent sat next to the infant and faced away from the displays. Parents were instructed to remain neutral and not to draw the infant's attention to or away from the displays.

A navy blue rectangular display board (74 cm × 30 cm) served as the background, about 60 cm from the infant. Displays consisting of black discs printed on white paper were glued on to smaller navy blue display boards measuring 52 cm × 21 cm, which were attached to and removed from the background display board between trials.

Habituation arrays consisted of 16 or 32 discs that varied in size and position within an array of a constant size (36 cm × 19 cm). The positions were chosen randomly from a 36 × 19 matrix and varied for each display. Displays were discarded if the disc positions looked nonrandom or cluttered. The overall size of the displays was twice as large those in Xu and Spelke (2000), and we could therefore use the same sizes for the individual discs (using the same display size would have required that individual discs be too small). The element density was .023 discs/cm^2 for the 16-element displays and .046 discs/cm^2 for the 32-element displays.

Over the habituation trials, the average area occupied by an individual element was twice as large for the 16-element arrays (mean element diameter = 1.83 cm, range = 1.06–2.37 cm) as for the 32-element arrays (mean element diameter = 1.30 cm, range = .75–1.67 cm), and so the average size of all the discs in an array combined and the average brightness of those discs were equated. At the infant's viewing distance of about 60 cm, each array subtended visual angles of 34° × 18°. Six arrays of 16 discs and six arrays of 32 discs were presented.

The test displays presented new arrays of 16 versus 32 discs. For the test arrays, the density of discs was equated, and therefore the 32-element arrays (38 cm × 24 cm) were twice as large as the 16-element arrays (24 cm × 19 cm). Moreover, the sizes of individual discs were equated (1.5 cm in diameter), and therefore the total size and average brightness in the 32-element arrays were twice those of the 16-element arrays. At a viewing distance of about 60 cm, the 16-element array subtended 22.5° × 18.0° and the 32-element array subtended 45° × 18°. The element density of the test arrays was constant at 0.035 element/cm^2 (see Figure 1 for a schematic depiction of the displays).[2]

Design

Half the infants were habituated to displays with 16 discs and half to displays with 32 discs. Equal numbers of males and females were assigned to each group. For each infant, the six habituation displays were presented in a random order. If the infant did not meet the habituation criterion after six trials, the displays were cycled in the same order until the end of the habituation sequence. Following habituation, infants were presented with six test trials in which displays with 16 discs and displays with 32 discs were shown alternately. The order of the test trials was counterbalanced across subjects.

Procedure

At the beginning of the test session, the experimenter used a squeaky toy to draw the infant's attention to the display board. She squeaked the toy at the top, bottom and the four far corners of the display to allow the observer to calibrate the infant's window of looking, and then the experiment began. On each trial, a curtain was raised to reveal a display, which remained present until the infant looked at it for at least 0.5 s and then looked away for 2 s continuously (or for a maximum look of 120 s), at which point the curtain was lowered for about 2 s to end the trial and allow the experimenter to change the display. Habituation trials continued until the infant either was given 14 trials or reached the habituation criterion of a 50% decline in looking time on three consecutive trials, relative to the total looking time on the first three trials that summed to at least 12 s. Fourteen of 16 infants reached the habituation criterion. After the habituation period, infants were shown six test trials following the same procedure as in the habituation phase.

Analyses

Because infants were habituated to a criterion and because all the test displays presented arrays that were novel on a variety of continuous dimensions, measures of the change in looking time from habituation to test are not reliable indexes of infants' responses to changes in number. Accordingly, all the analyses focused on the looking patterns exhibited during the test.[3] Test trial looking times were edited to remove all outlying scores (> 2 SD from the mean for each condition): six test trials were replaced by condition means because they were outliers. Looking times then were subjected to a 2 × 3 × 2 mixed-factor analysis of variance (ANOVA) testing the between-subject factor of Habituation condition (16 or 32) and the within-subject factors of Test trial pair (first, second or third) and Test trial type (old or new number). In addition, looking preferences were subjected to nonparametric analyses, and looking times in the present experiment were compared to those obtained in past experiments using the same method (Xu & Spelke, 2000) by further ANOVAs with the additional between-subject factor of Experiment.

Results

Figure 2 presents the mean looking times during the first three and last three habituation trials and during the three test trials presenting the old versus new number. During the test, infants looked longer at the new number (M = 8.9 s) than at the old number (M = 6.5 s). The ANOVA revealed a main effect of Test trial type, $F(1, 14) = 6.398$, $p < .05$, and no other statistically significant effects. Twelve of the 16 infants (11 of the 14 infants who reached habituation criterion) looked longer at the new-number displays (Wilcoxin $z = 2.534$, $p < .01$).

Further analyses compared infants' performance discriminating 16- from 32-element displays in Experiment 1

---

[2] Following standard practice (Durgin & Proffitt, 1996; Beaudot & Mullen, 2000), we defined density as number of elements per unit area and equated the distance between habituation and test displays based on this measure of density. However, it is possible that the test displays for 32 elements would be 'clumpier' than those for 16 elements, because if all the elements were randomly placed on the displays, the more elements there are, the more likely that some of them would cluster together. We tested this possibility empirically by calculating the distance of each element from its nearest neighbor, and averaging these distances for each display. The average distances for the 16-element displays were 2.84, 3.62 and 3.68 (grand average 3.38) and those for the 32-element displays were 3.18, 2.98 and 3.56 (grand average 3.24). Thus the displays were equally clustered.

[3] Although it is common to compare looking time during habituation to looking time during test, using 'dishabituation' as a measure of discrimination and categorization, the design of the present experiments excludes such comparisons. In the present experiments, pairs of linked continuous variables such as array size and density are de-confounded from number by equating one variable in the habituation arrays at the two numerosities and equating the other variable in the test arrays at the two numerosities. By design, therefore, the continuous variables that must be de-confounded from number are equated within habituation arrays and within test arrays, but not across habituation and test arrays: The last habituation array and the first novel-number test array differ not only in number but also in item size, item density, array size, and summed area. Because any measure of dishabituation would confound changes in these variables with changes in number, only analyses comparing performance on the same-number and different-number test trials serve to assess discrimination of numerosity.

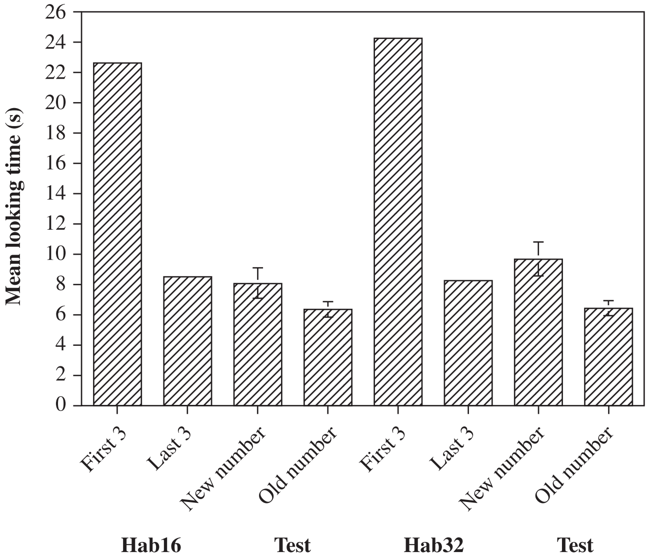

Figure 2 *Mean looking times and standard errors during the first three and last three habituation trials and during the six test trials of Experiment 1.*

to infants' performance discriminating eight- from 16-element displays in Xu and Spelke's (2000) research. A 2 (Experiment) × 2 (Test trial type) ANOVA revealed significant effects of Experiment, $F(1, 30) = 6.342$, $p < .05$, and of Test trial type, $F(1, 30) = 11.297$, $p < .005$, and more importantly, no interaction between these factors. Infants looked longer overall at the larger displays of Experiment 1 than at the smaller displays used by Xu and Spelke (2000), and infants in both experiments looked longer at the new number on test trials. The latter tendency was equally strong in the two studies.

Discussion

The findings of Experiment 1 replicate and extend those of Xu and Spelke (2000) and Xu (2003b). They provide further evidence that 6-month-old infants discriminate between visual arrays with large numbers of elements on the basis of numerosity. Infants' numerosity discrimination extends to set sizes that are twice as large as any yet investigated: 16 versus 32. The next experiment therefore investigates whether infants' numerosity discrimination, like that of adults, is limited by the ratio difference between the set sizes.

Experiment 2

In past experiments, infants successfully discriminated arrays of eight versus 16 discs and they failed to discriminate arrays of 8 versus 12 discs (Xu & Spelke, 2000). Given that the infants in Experiment 1 discriminated between visual arrays with 16 versus 32 discs, it is possible to test whether infants' discrimination is subject to a set-size ratio limit for numerosities in this larger range by presenting them with a discrimination task with 16 versus 24 disc arrays. If infants' numerosity discrimination is subject to a ratio limit between 1.5 and 2.0, as suggested by experiments testing discrimination of smaller numerosities, then infants should fail to discriminate 16- from 24-element arrays. In contrast, if discrimination depends on the absolute difference in set sizes, then infants should succeed at this discrimination.

Method

The method was the same as in Experiment 1, except as follows. Participants were 16 infants (eight males and eight females) ranging in age from 5 months, 15 days to 6 months, 10 days (mean age = 5 months, 22 days). An additional three infants were eliminated from the sample because of fussiness. The displays for Experiment 2 were created by doubling the size of the displays used in Xu and Spelke's (2000) study of infants' discrimination of eight versus 12 discs and presenting the same individual disc sizes and arrangements as in that study. The size of the habituation displays was 38 cm × 20 cm. The average size of the individual discs in the 16-element habituation displays was 2.24 cm, ranging from 1.41 cm to 2.82 cm, and the average density was .021 element/cm^2. The average size of individual discs in the 24-element displays, therefore, was 1.82 cm, ranging from 1.14 cm to 2.44 cm, and the average density was .031 element/cm^2. All the test arrays displayed discs of the same size (2 cm in diameter) and density (0.026 element/cm^2). The image size was 32 cm by 19 cm for the 16-disc test displays and 38 cm by 24 cm for the 24-disc displays. As in Experiment 1, therefore, the overall image size and total filled area were equated for the two numerosities during habituation, and the element sizes and densities were equated for the two numerosities during the test. Interobserver reliability again averaged 92%. Twelve of 16 infants reached the habituation criterion. For the analyses, four test trials were replaced by condition means because they were outliers.

Results

Figure 3 presents the mean looking times during the habituation and test trials. Infants did not look longer at the test displays presenting a novel numerosity than at the displays presenting the familiar numerosity. The 2 (Habituation condition: 16 versus 24) × 3 (Test trial pair)

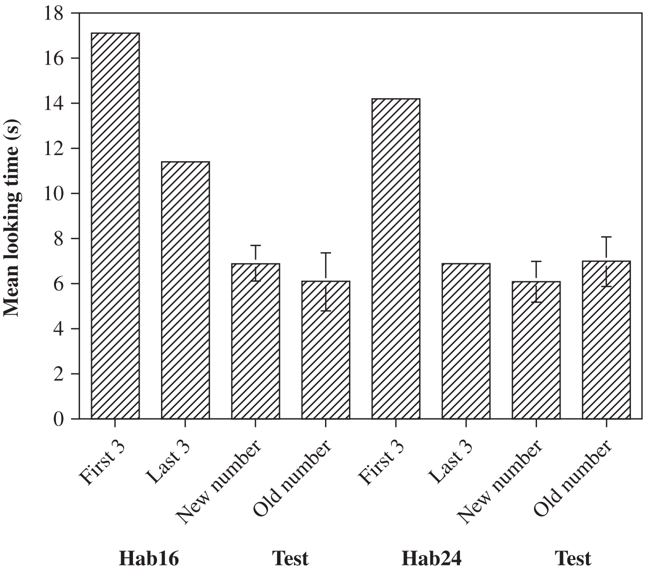

Figure 3 *Mean looking times and standard errors during the first three and last three habituation trials and during the six test trials of Experiment 2.*

× 2 (Test trial type: old versus new number) ANOVA revealed no significant effects. In particular, there was no main effect of Test trial type ($F < 1$). Infants looked equally at new number displays ($M = 6.4$ s) and old number displays ($M = 6.5$ s). Seven of the 16 infants (six of the 12 infants who reached habituation criterion) looked longer at the test displays with the novel numerosity (Wilcoxin $z < 1$).

A further analysis compared looking patterns in Experiment 2 to those in Experiment 1. The 2 (Experiment) × 2 (Test trial type) ANOVA revealed a significant interaction of Experiment by Test trial type, $F(1, 30) = 4.069$, $p < .05$, and no other effects. The preference for the novel numerosity was greater for the task of discriminating 16 from 32 than for the task of discriminating 16 from 24.

Discussion

Although infants successfully discriminated 16- from 32-element arrays, they failed to discriminate 16- from 24-element arrays when tested with the same method and type of displays. Performance was significantly worse in the present experiment than in Experiment 1, providing evidence that infants show different discrimination abilities when presented with sets in different ratios. These findings complement the findings of Xu and Spelke (2000), which revealed successful discrimination of eight from 16 and failure to discriminate eight from 12. Together, the findings provide evidence that infants' performance on a numerosity discrimination task depends on the ratio difference in set sizes: infants discriminate between sets that differ by a ratio of 2.0 but not 1.5. Experiments 1 and 2, therefore, reveal a qualitative similarity between numerosity discrimination in human infants, human adults and non-human animals: in all these populations, numerosity discrimination succeeds over a wide numerical range and is subject to a set size ratio limit.

Experiment 3

Although the infants in Experiment 1 and in past research discriminated between large numbers of visual forms when a variety of continuous variables were controlled, infants have shown no evidence of discriminating between small numbers of visual forms when the number of forms is inversely correlated with the total contour length in the display (Clearfield & Mix, 1999). Three differences between the present research and that of Clearfield and Mix could account for these discrepant findings: the different numerical magnitudes tested (large numbers versus small numbers), the different continuous quantity controls (total filled area versus total contour length), or a host of differences in method.[4] In Experiment 3, we begin to distinguish these possibilities by investigating infants' discrimination between arrays of 16 versus 32 elements under conditions that control for total contour length. Infants were habituated and tested with arrays of 16 versus 32 elements whose sizes were chosen so that the total contour length presented for the two numerosities was equated during the habituation sequence, whereas item size and density were equated during test. When the total contour length of displays differing in numerosity is equated, total area is correlated negatively with number. If infants respond to number and not to either total contour length or summed area, therefore, they should look longer at the test displays with the novel numerosity. If infants respond to contour length only, they should respond equally to test displays at the two numerosities. If infants respond to summed area, they should look longer at the test displays with the familiar numerosity, since those displays differ more in summed area from the habituation displays.

---

[4] Perhaps the most important difference in method between the two sets of experiments is that the studies of Xu and Spelke (2000) and of Experiments 1 and 2 equated for continuous variables across the two numerosities, whereas the studies of Clearfield and Mix (1999) presented continuous variables that were inversely correlated with number: the test display with the novel numerosity presented a familiar total contour length, and the test display with the familiar numerosity presented a novel total contour length. Thus, Clearfield and Mix's studies leave open the possibility that infants detected both the change in contour length and the change in numerosity, and that the former change had a greater influence on their looking time.

Method

The method was the same as in Experiment 1, except as follows. Participants were 16 full-term infants (eight males, eight females) ranging in age from 5 months, 17 days to 6 months, 15 days (mean age = 6 months, 2 days). An additional three infants were eliminated from the sample because of fussiness. The habituation arrays were the same as in Experiment 1, except for the sizes of the discs. The average circumference of each individual disc was twice as large for the 16-disc arrays (mean element diameter = 1.77 cm, range = 1.06–2.36 cm) as for the 32-disc arrays (mean element diameter = .88 cm, range = .53–1.18 cm). Habituation displays at the two numerosities therefore were equated for display size and total contour length, and they differed in item size and density. The test arrays were the same as in Experiment 1, except for the size of the individual discs: all discs were 1.33 cm in diameter. Test displays at the two numerosities therefore were equated for item size and density and differed in display size and total contour length. Interobserver reliability averaged 90%. Eleven of the 16 infants reached the habituation criterion. For the analyses, four test trials were replaced by condition means because they were outliers.

Results

Figure 4 presents the mean looking times during the habituation and test trials. During the test, infants looked longer at the novel numerosity ($M = 12.8$ s) than at the familiar numerosity ($M = 8.8$ s). The ANOVA testing for effects of Habituation condition (16 or 32), Test trial pair, and Test trial type (old or new) revealed a main effect of Test trial type, $F(1, 14) = 7.10$, $p < .05$, and no other significant main effects or interactions. Thirteen of the 16 infants (10 of the 11 infants who reached habituation criterion) looked longer at the novel numerosity (Wilcoxin $z = 3.103$, $p < .01$).

An additional ANOVA compared the present experiment with Experiment 1, with Experiment and Test trial type as factors. The ANOVA revealed a main effect of Test trial type, $F(1,30) = 13.051$, $p < .001$, and no other statistically significant effects. In both experiments, infants looked longer at the new numerosity.

Discussion

Six-month-old infants successfully discriminated between arrays of 16 versus 32 discs when the arrays were controlled for contour length, density, image size and item size, and when summed area correlated negatively rather than positively with number. This finding replicates and extends that of Experiment 1 and of previous research probing infants' discrimination of large numerosities. It provides evidence that infants' numerosity discrimination is robust over variation in a range of continuous variables, including element size and density, image size, total filled area, total contour length and a host of perceptual properties that depend on these variables including surface texture, brightness and contrast.

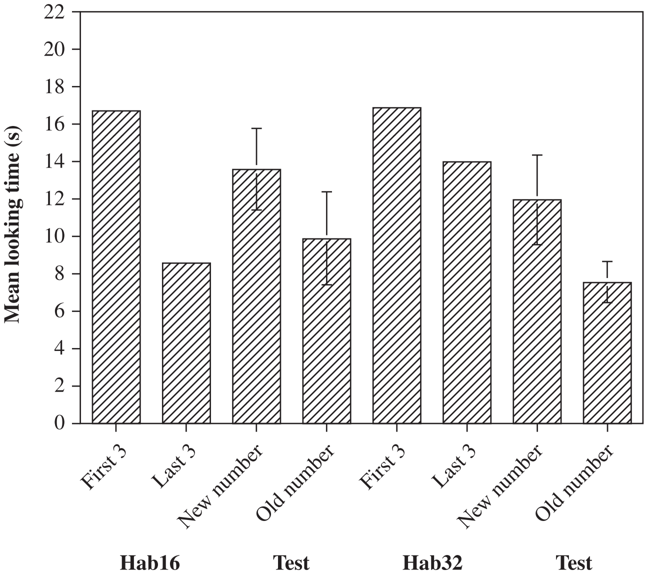

Figure 4 *Mean looking times and standard errors during the first three and last three habituation trials and during the six test trials of Experiment 3.*

The results of Experiment 3 indicate that the divergent findings of Xu and Spelke (2000 and Experiment 1) and of Clearfield and Mix (1999) do not stem from differences in continuous-quantity controls. Like Clearfield and Mix's studies, Experiment 3 systematically varied the total contour length presented in the displays. In contrast to their studies, however, infants were found to discriminate numerosities despite this variation.[5]

---

[5] Mix, Clearfield and Drake (under review) tested 6-month-old's discrimination of large-number displays in which number and continuous variables were inversely correlated. After habituation to one array of eight or 16 discs of the same unchanging size, infants were tested with new arrays of eight versus 16 discs in which the items in the less numerous array were eight times as large as those in the more numerous array. Thus, the test array with the novel numerosity presented the same total contour length as the habituation array, whereas the test array with the familiar numerosity presented a dramatic change in all continuous variables. Under these conditions, infants responded more to the change in continuous variables than to the change in number. This finding provides evidence that infants respond to an eight-fold change in contour length more than to a two-fold change in number, but it is fully compatible with the evidence of Xu (2003b), Xu and Spelke (2000), and of the present studies that infants are sensitive to changes in numerosity when continuous variables are controlled (see also Brannon *et al.*, in press).

Similarly, Xu (2003b) found successful discrimination with four versus eight discs when total contour length was controlled for. In the final experiment, we ask whether the differing findings from the two sets of studies stem from differences in the sizes of the numerosities presented for discrimination (large versus small), or from other differences in their methods, by using the method of Xu and Spelke (2000) to investigate infants' discrimination of small numbers of elements.

Experiment 4

Experiment 4 tested infants' discrimination between visual arrays containing one versus two elements. As in our previous experiments, the elements were discs whose sizes and distribution were varied systematically so as to control for a variety of continuous variables. In order for Experiment 4 to be as comparable as possible to past research with large numerosities, each display containing one disc was matched in display size and area to one of the displays containing eight discs in Xu and Spelke's (2000) first experiment, and each display containing two discs was matched in display size and total filled area to one of the displays containing 16 discs in that experiment. The individual visual forms presented in Experiment 4 therefore were eight times as large as those presented in Experiment 1, and the full displays were smaller than those in Experiment 1 (and equal to those of Xu and Spelke's first experiment). Pilot testing suggested that these displays would evoke levels of visual attention that were similar to those in Xu and Spelke's (2000) experiments, whereas displays containing one or two discs of the same dimensions as those used in Experiment 1 evoked shorter looking times.

Method

Participants were 16 infants (eight males, eight females) ranging in age from 5 months, 15 days to 6 months, 15 days (mean age = 6 months, 2 days). An additional six infants were eliminated from the sample because of fussiness or parental interference. The method was the same as in our past experiments, except as follows. The one-element habituation displays consisted of a single, randomly placed disc; the two-element habituation displays consisted of two such discs. All the arrays presented discs with eight times the area of the displays of Experiment 1, and they occupied a reduced range of spatial positions chosen to accommodate discs of these sizes. As in past studies, therefore, the image sizes and total filled area were equated across the two numerosities for the habituation phase, and the element sizes and densities of the arrays were equated across the numerosities for the test phase. Moreover, the total filled area and contour length of the present arrays were identical to those of the arrays in Xu and Spelke's (2000) study of discrimination of eight from 16 discs. The single form in the one-element displays varied in size from 3.0 cm to 6.7 cm in diameter during familiarization (mean diameter = 5.16 cm) and was 4.24 cm in diameter during test, and each form in the two-element displays had half the area of its counterpart in the one-element displays during habituation and the same area as its counterpart during the test (Figure 5). Interobserver agreement averaged 90%. Ten out of 16 infants reached habituation criterion. Six test trials were replaced by condition means because they were outliers.

Results

Figure 6 presents the mean looking times on the habituation and test trials. Infants looked about equally at the test displays presenting a novel ($M = 5.2$ s) versus familiar numerosity ($M = 5.1$ s). The 2 (Habituation condition) × 3 (Trial pair) × 2 (Test trial type: old versus new numerosity) ANOVA revealed no significant effects, including no main effect of Test trial type, $F(1, 14) < 1$. Seven of the 16 infants (six of the 10 infants who reached habituation criterion) looked longer at the test displays with the novel number (Wilcoxin $z < 1$).

Further analyses compared the test trial looking patterns of the infants in Experiment 4 to those of their counterparts in Experiment 1, who were tested with large numerosities in the same difference ratio and with displays that were similar in total filled area and image size. This 2 × 3 × 2 ANOVA revealed a significant main effect of Experiment, $F(1, 30) = 15.35$, $p < .001$, and a significant interaction of Experiment by Test display type, $F(1, 30) = 4.57$, $p < .05$. Infants showed higher looking overall when presented with the large number displays, which covered a larger area than the small number displays. Most important, infants showed a reliable preference for the novel numerosity when given the task of discriminating 16 from 32 discs but not when given the task of discriminating one from two discs. A final analysis compared the test trial looking patterns of the infants in Experiment 4 to those in Experiment 1 of Xu and Spelke (2000), who viewed displays of eight versus 16 discs that exactly matched the present displays in overall size and in continuous extent. This ANOVA revealed no main effect of Experiment ($F < 1$): infants looked just as long at the displays of small numbers of large discs (mean, 5.0 s) as at the displays of large numbers of small discs (mean, 5.5 s). Nevertheless, there was a borderline-significant interaction of Experiment by

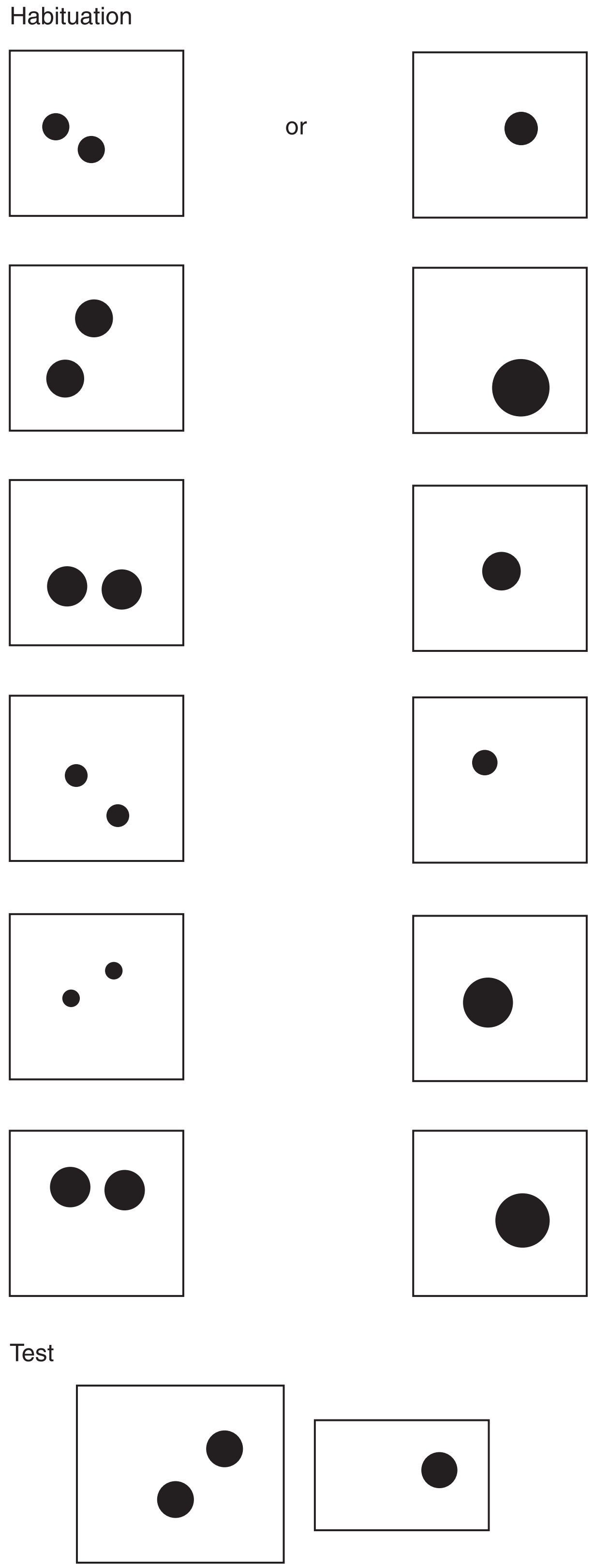

Figure 5 *Selected habituation and test displays for Experiment 4.*

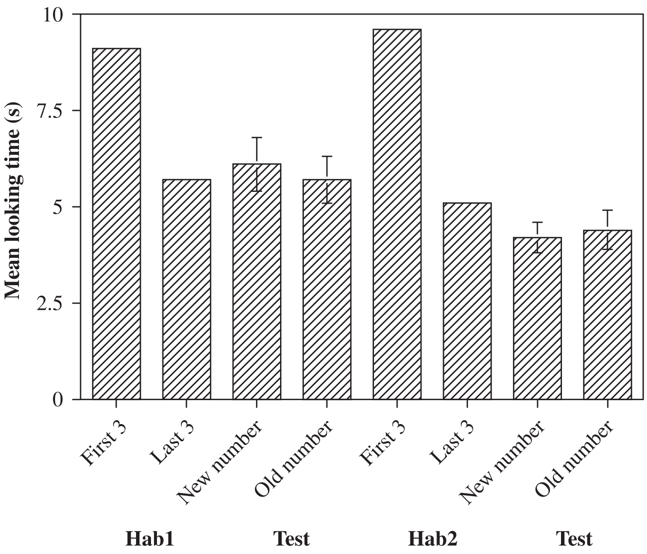

Figure 6 *Mean looking times and standard errors during the first three and last three habituation trials and during the six test trials of Experiment 4.*

Test display type, $F(1, 30) = 3.87$, $p < .06$. Infants showed a marginally greater preference for the novel number display when tested with large (eight versus 16) than with small (one versus two) numerosities.

Discussion

Experiment 4 provides no evidence that infants discriminate between arrays of small numbers of visual elements on the basis of numerosity, when continuous variables of total filled area, image size, item size, and item density are controlled. This finding replicates and extends the findings of Clearfield and Mix (1999), Feigenson, Carey and Spelke (2002), and Xu (2003b), who showed that infants fail to discriminate between small numbers of visual forms or objects when the continuous variables of total contour length, total filled area or total volume are controlled. The findings call further into question the evidence for small-number discrimination in infancy from studies that confound numerical differences with differences in these continuous variables, and they are consistent with theories that attribute small-number discrimination abilities to mechanisms for representing continuous quantities (Clearfield & Mix, 1999; Mix, Huttenlocher & Levine, 2002) or mechanisms for representing individual objects (Simon, 1997; Carey, 2001).

The infants in Experiment 4, who were presented with the task of discriminating one from two discs, showed reliably less discrimination than those in Experiment 1, who were presented with the task of discriminating 16 from 32 discs, and marginally less discrimination than those in Xu and Spelke's (2000) first experiment, presented with the task of discriminating eight from 16 discs. These differing findings are striking, because the different discriminations were tested on the same population of infants, by means of the same method and with closely matched displays. In particular, the displays used in Experiment 4 presented the same type of elements with the same continuous quantity controls as in our previous studies, they tested discrimination at the same difference ratio that yielded successful discrimination in past studies with large numerosities, and they used displays that evoked looking times that were as long as those in Xu and Spelke's (2000) experiments (although shorter than those of Experiment 1, which presented displays that covered a larger area). Infants' significantly poorer performance with the small-number displays therefore suggests that their sense of number is not evoked by arrays with very small numerosities, at least under the conditions of the present experiments.

General discussion

Four experiments investigated number discrimination in 6-month-old infants under conditions that controlled for continuous variables that are typically correlated with number. Experiments 1 and 3 replicate Xu (2003b) and Xu and Spelke's (2000) findings that infants represent large numerosities under conditions that control for contour length as well as total filled area, display size, element size, element density and correlated perceptual variables such as surface brightness, contrast and texture. Moreover, Experiments 1 and 3 extend the evidence for large number discrimination to numerosities that are larger than any previously tested: 16 versus 32. The positive findings of these experiments suggest that infants discriminate large numerosities robustly, over a considerable range of numerosities and over changes in every continuous variable that has been considered, to our knowledge, in the literature on numerical discrimination of two-dimensional visual-spatial arrays.

What are the mechanisms by which infants discriminate numerosities, and how do they compare to the mechanisms of numerosity discrimination by older children and adults? Building on the findings of Lipton and Spelke (2003), Xu and Spelke (2000), and Xu (2003b), Experiments 1 and 2 provide evidence that infants' discrimination of large numerosities is subject to a set size ratio limit, with successful discrimination when numerosities differ by a ratio of 2.0 (eight versus four, 16 versus eight and 32 versus 16) and failure when numerosities differ by a ratio of 1.5 (six versus four, 12 versus eight and 24 versus 16). Because the set size ratio limit is a signature of the mechanisms of large number representation in human adults who are prevented from counting (Cordes *et al.*, 2001; Van Oeffelen & Vos, 1982; Whalen, Gelman & Gallistel, 1999), in human children, who cannot count large sets reliably (Huntley-Fenner, 2001), and in other animals (Gallistel, 1990; Hauser, Tsao, Garcia & Spelke, 2003), the present findings add empirical support to the thesis that a common system of number representation, shared by humans and other animals, is present and functional in 6-month-old infants.

Further evidence for a common mechanism of number representation in infants and adults comes from studies of infants' discrimination of the number of sounds in an auditory sequence. As already noted, 6-month-old infants discriminate sequences of eight sounds from sequences of 16, but not 12, sounds when the continuous variables of sound duration and amount, sequence duration and sound rate are controlled (Lipton & Spelke, 2003). These findings are consistent with the thesis that a similar ratio limit on numerical discrimination applies to visual-spatial arrays and to auditory-temporal arrays for human infants, as it does for human adults (Barth, Kanwisher & Spelke, 2003). This possibility was further confirmed recently by studies that tested for the same ratio limit with different numbers of sounds: 6-month-old infants discriminated sequences of four versus eight sounds (a 2.0 ratio) but not four versus six sounds (a 1.5 ratio) (Lipton & Spelke, in press).

Finally, Experiments 3 and 4 shed further light on a conflict in the literature on infants' numerosity discrimination: Although infants discriminate between large numerosities when continuous variables are controlled (e.g. Brannon, 2002; Lipton & Spelke, 2003; Xu, 2003b; Xu & Spelke, 2000), they have failed to discriminate between small numerosities when tested with displays in which continuous variables were controlled or inversely correlated with number (Clearfield & Mix, 1999; Feigenson, Carey & Spelke, 2002; Xu, 2003b). These results, along with Xu (2003b), provided evidence that the divergent findings stem from differences in the range of numerosities tested, independently of any differences in the methods or stimulus controls used in the different studies. The infants in Experiment 3 successfully discriminated between large numerosities (16 versus 32), even though the experiment controlled for the same continuous variable – total contour length – as the studies of Clearfield and Mix (1999) that showed no evidence for discrimination of small numerosities. Moreover, the infants in Experiment 4 failed to discriminate between small numerosities (one versus two) when tested with the same method and stimulus controls as in Experiment 1 and the studies of Xu (2003b) and Xu and Spelke (2000), which showed successful discrimination of large numerosities that differed by the same ratio. Infants' performance was reliably worse in Experiment 4 than in Experiment 1, even though the two experiments used the same design and method and closely similar displays. Together, these experiments suggest that infants discriminate between large numerosities better than between small numerosities, at least under the conditions of the present studies.

Our findings raise several questions. First, what are the mechanisms of large number discrimination? One hypothesized mechanism is an iterative, counting-like mechanism that tags each item in an array successively and sums activation from all the elements (e.g. Gallistel & Gelman, 2000). A second hypothesized mechanism is similar but non-iterative: it tags every item in an array in parallel and sums activation over the items (e.g. Dehaene & Changeux, 1993). A third hypothesized mechanism computes number by assessing the area and density of an array of elements and multiplying these values (e.g. Barth *et al.*, 2003; Church & Broadbent, 1990). Studies of adults and of infants cast doubt on the first hypothesis, because the time to discriminate numerosities in dot arrays is independent of numerical magnitude (Barth *et al.*, 2003). To date, no evidence distinguishes the second and third hypotheses, to our knowledge.

A second question concerns developmental changes in large number discrimination: why is large-number discrimination so imprecise for infants, and how does its precision improve with development? There is a considerable quantitative difference between the performance of infants in large number discrimination experiments and the performance of human adults and adult non-human animals. Whereas human infants discriminate between numerosities with a ratio difference of 2.0 but not 1.5, human adults who are instructed to attend to numerosity discriminate between numerosities with a ratio difference as small as 1.15 (e.g. Barth *et al.*, 2003; Van Oeffelen & Vos, 1982). Adult non-human primates also show high sensitivity to numerosity differences when they are trained to respond to number (e.g. Brannon & Terrace, 1998; Matsuzawa, 1985; Mechner, 1958; Nieder, Freedman & Miller, 2002). Although the spontaneous number discrimination abilities of monkeys are lower than those of trained animals, untrained monkeys discriminate between numerosities with a ratio difference between 1.25 and 1.5 (Hauser *et al.*, 2003), outperforming 6-month-old infants tested with similar displays and methods (Lipton & Spelke, 2003).

Although the reasons for the imprecision of infants' large-number representations are not clear, recent research suggests that these representations increase in precision over the infancy period. Nine-month-old infants, tested with the same auditory sequences and head-turn preference method as 6-month-old infants, successfully discriminated sequences of eight sounds from sequences of 12 sounds but not eight from 10 sounds, providing evidence that the ratio limit on number discrimination declined from 2.0 to 1.5 between 6 and 9 months, at least for these displays (Lipton & Spelke, 2003). Similarly, 10-month-old infants succeeded in discriminating eight discs from 12 discs but not eight from 10 discs (Xu & Arriaga, 2004). These findings suggest that the developmental increase in sensitivity to numerical differences does not stem from the acquisition of formal arithmetic or verbal counting, but rather from processes of neural maturation and perceptual experience that occur already in the first year of life. It remains possible, however, that experience with verbal counting or formal training in arithmetic further sharpens children's sensitivity to numerosity.

A third question concerns infants' failure to discriminate between the smallest numbers of visual elements. Why do infants fail to discriminate one versus two objects or visual forms when continuous spatial quantities are controlled? Hints of an answer may come from the literature on numerical estimation in human adults, which distinguishes between two systems of numerical representation: an exact system that operates in the small number range (the 'subitizing' or 'object tracking' system, e.g. Trick & Pylyshyn, 1994) and an approximate system (e.g. Dehaene, 1997). More specifically, infants and adults may represent information about small numbers of elements by means of a system that permits parallel, attentive tracking of up to three or four objects (see Leslie, Xu, Tremoulet & Scholl, 1998; Simon, 1997; Xu, 1999; and Xu, 2003a, for detailed discussion of object tracking in infancy; see Kahneman, Treisman & Gibbs, 1992; Scholl, 2001; Scholl & Pylyshyn, 1999; and Trick & Pylyshyn, 1994, for relevant evidence from adults). This system serves to track individual, persisting objects as they move through a scene, but it does not serve explicitly to enumerate objects or perform numerical comparisons. When the infants in Experiment 4 were presented with the succession of habituation displays, for example, this system would register different individual objects on every trial (e.g. objects A and B on trial 1, C and D on trial 2, E and F on trial 3, etc.), but would fail to register that the *number* of objects presented on successive trials was the same (e.g. *two* objects on trial 1, *two* objects on trial 2, etc.). The second system represents approximate numerosities and underlies large number discrimination in adults, non-human animals and infants (Dehaene, 1997; Gallistel, 1990). The signature properties that distinguish the small exact and large approximate systems are their differing limits: the set-size limit on the small-number system (up to four for adult humans and non-human primates, up to three for infants) and the set-size ratio limit on the approximate system (as low as 1.15 for adults, between 1.5 and 2.0 for 6-month-old infants).

Why does the system of number sense fail to apply to small numerosities, such that infants fail to discriminate between arrays of one versus two visual elements in studies that control for continuous variables? Three possible answers to this question appear to be worthy of further investigation (see also Xu, 2003b). First, the system for representing approximate numerosities may fail to operate on the smallest numbers because its computations are unstable or undefined for small values. For example, if human adults and non-human animals estimate approximate numerosities in visual-spatial arrays by assessing the area and density of an array of elements and multiplying these values (e.g. Church & Broadbent, 1990), their estimation process necessarily will fail for the smallest numerosities: inter-item distance is undefined for a one-object array, area is undefined for arrays of less than three objects, and neither measure is a reliable index of numerosity when small numbers of items are randomly distributed in an array. The approximate number system, therefore, may be used only to estimate larger numbers of items. Second, the system for representing approximate numerosities may operate on sets of all sizes, but its outputs may be inhibited by the operation of the object tracking system. For example, infants presented with two objects may focus attention on the individual objects, not on the numerosity of the set that they compose, and so they may fail to register the set's cardinal value. Third, the system for representing approximate numerosities may operate on sets of all sizes, but the representations that it delivers may fail to evoke novelty preferences when small numbers of items appear with the present displays and methods. Converging studies of adults and infants could serve to investigate each of these possibilities.

However these questions are answered, the signature limits revealed by the present experiments provide evidence for continuity in core numerical representations, both over human ontogeny and over primate phylogeny. These core representations may provide the building blocks upon which children construct the uniquely human, natural number representations that underpin children's verbal counting and symbolic arithmetic.

Acknowledgements

Supported by NSF grant (BCS-0004235) to Fei Xu and by NIH grant (R37-HD23102) to Elizabeth S. Spelke. We thank Stanislas Dehaene, Ghislaine Dehaene-Lambertz, Jennifer Lipton, Cristina Sorrentino, and three anonymous reviewers for helpful discussions and comments on an earlier draft of the paper. We also thank members of the Northeastern and MIT Infant Cognition laboratories for assistance in data collection, and the parents and the infants for their participation.

References

Antell, S.E., & Keating, L.E. (1983). Perception of numerical invariance by neonates. *Child Development*, **54**, 695–701.

Balakrishnan, J.D., & Ashby, F.G. (1992). Subitizing: magical numbers or mere superstition? *Psychological Research*, **54**, 80–90.

Barth, H., Kanwisher, N., & Spelke, E. (2003). Construction of large number representations in adults. *Cognition*, **86**, 201–221.

Beaudot, W.H.A., & Mullen, K.T. (2000). Role of chromaticity, contrast, and local orientation cues in the perception of density. *Perception*, **29**, 581–600.

Bijeljac-Babic, R., Bertoncini, J., & Mehler, J. (1991). How do four-day-old infants categorize multisyllabic utterances? *Developmental Psychology*, **29**, 711–721.

Brannon, E.M. (2002). The development of ordinal numerical knowledge in infancy. *Cognition*, **83**, 223–240.

Brannon, E.M., & Terrace, H.S. (1998). Ordering of the numerosities 1 to 9 by monkeys. *Science*, **282**, 746–749.

Brannon, E.M., Abbot, S., & Lutz, D.J. (in press). Number bias for the discrimination of large visual sets in infancy. *Cognition*.

Carey, S. (2001). Evolutionary and ontogenetic foundations of arithmetic. *Mind and Language*, **16**, 37–55.

Church, R.M., & Broadbent, H.A. (1990). Alternative representations of time, number and rate. *Cognition*, **37**, 55–81.

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

Cordes, S., Gelman, R., Gallistel, C.R., & Whalen, J. (2001). Variability signatures distinguish verbal from nonverbal counting for both large and small numbers. *Psychonomic Bulletin and Review*, **8** (4), 698–707.

Dehaene, S. (1997). *The number sense*. New York: Oxford University Press.

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

Durgin, F.H., & Proffitt, D.R. (1996). Visual learning in the perception of texture: simple and contingent aftereffects of texture density. *Spatial vision*, **9**, 423–474.

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

Feigenson, L., Carey, S., & Spelke, E.S. (2002). Infants' discrimination of number vs. continuous extent. *Cognitive Psychology*, **44**, 33–66.

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

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

Hauser, M.D., Tsao, F., Garcia, P., & Spelke, E.S. (2003). Evolutionary foundations of number: spontaneous representation of numerical magnitudes by cotton-top tamarins. *Proceedings of the Royal Society, London*, **B 270**, 1441–1446.

Huntley-Fenner, G. (2001). Children's understanding of number is similar to adults' and rats': numerical estimation by 5- to 7-year-olds. *Cognition*, **78**, B27–40.

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

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

Lipton, J.S. (2003). Preschool children's mapping of number words to non-symbolic numerosities. Poster presented at the biennial meeting of the Society for Research in Child Development, Tampa, USA.

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

Lipton, J.S., & Spelke, E.S. (in press). Discrimination of large and small numerosities by human infants. *Infancy*.

Mandler, G., & Shebo, B.J. (1982). Subitizing: an analysis of its component processes. *Journal of Experimental Psychology: General*, **11**, 1–22.

Matsuzawa, T. (1985). Use of numbers by a chimpanzee. *Nature*, **315**, 57–59.

Mechner, F. (1958). Probability relations within response sequences under ratio reinforcement. *Journal of the Experimental Analysis of Behavior*, **1**, 109–121.

Mix, K., Huttenlocher, J., & Levine, S. (2002). Multiple cues for quantification in infancy: is number one of them? *Psychological Bulletin*, **128**, 278–294.

Mix, K.S., Clearfield, M.W., & Drake, P.E. (under review). Infants' use of amount to discriminate large sets. Manuscript submitted for publication.

Nieder, A., Freedman, D.J., & Miller, E.K. (2002). Representation of the quantity of visual items in the primate prefrontal cortex. *Science*, **297**, 1708–1711.

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.

Sharon, T., & Wynn, K. (1998). Infants' individuation of actions from continuous motion. *Psychological Science*, **9**, 357–362.

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

Spelke, E.S. (2000). Core knowledge. *American Psychologist*, **55**, 1233–1243.

Starkey, P., & Cooper, R.G. (1980). Perception of numbers by human infants. *Science*, **210**, 1033–1035.

Starkey, P., Spelke, E.S., & Gelman, R. (1983). Detection of intermodal numerical correspondences by human infants. *Science*, **222**, 179–181.

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

Temple, E., & Posner, M.I. (1998). Brain mechanisms of quantity are similar in 5-year-old children and adults. *Proceedings of the National Academy of Sciences*, **95**, 7836–7841.

Treiber, F., & Wilcox, S. (1984). Discrimination of number by infants. *Infant Behavior and Development*, **7**, 93–100.

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.

Van Oeffelen, M.P., & Vos, P.G. (1982). A probabilistic model for the discrimination of visual number. *Perception & Psychophysics*, **32**, 163–170.

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. (2002). Enumeration of collective entities by 5-month-old infants. *Cognition*, **83**, B55–B62.

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. (2003a). The development of object individuation in infancy. In J. Fagen & H. Hayne (Eds.), *Progress in infancy research*, Vol. 3 (pp. 159–192). Mahwah: Lawrence Erlbaum.

Xu, F. (2003b). Numerosity discrimination in infants: evidence for two systems of representations. *Cognition*, **89**, B15–B25.

Xu, F., & Arriaga, R. (2004). Large number discrimination in 10-month-old infants. Manuscript in preparation.

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

Received: 1 October 2003
Accepted: 20 April 2004