INFANCY, *9*(2), 173–189

The Effect of Heterogeneity on Numerical Ordering in Rhesus Monkeys

Jessica F. Cantlon and Elizabeth M. Brannon
Center for Cognitive Neuroscience and
Department of Psychological & Brain Sciences
Duke University

We investigated how within-stimulus heterogeneity affects the ability of rhesus monkeys to order pairs of the numerosities 1 through 9. Two rhesus monkeys were tested in a touch screen task where the variability of elements within each visual array was systematically varied by allowing elements to vary in color, size, shape, or any combination of these dimensions. We found no evidence of a cost (or benefit) in accuracy or reaction time when monkeys were tested with stimuli that were heterogeneous in color, size, or shape. This was true even though both monkeys experienced extended initial training with arrays that were homogeneous in the color, shape, and size of elements. The implications of this finding for the mechanisms that monkeys use to represent and compare numerosities are discussed.

Adult humans possess rich and profoundly abstract number concepts. We recognize the numerical equivalence between sets as diverse as four painters and four strokes of a brush, three musicians and three saxophone notes, or two books and two ideas. Such examples illustrate that, as adult humans, we form numerical representations when perceiving simultaneously occurring visual stimuli (e.g., two books), successively occurring visual events (e.g., strokes of a brush), successively occurring auditory events (e.g., three saxophone notes), or when thinking about sets as intangible as an author's ideas. In addition to abstracting number across varying stimulus formats (e.g., successive vs. simultaneously occurring objects or events) and across different stimulus modalities (e.g., auditory or visual), adult human number representations are also abstract in that they are independent of perceptual aspects of the stimuli (e.g., size, shape, color). For example, two

Correspondence should be addressed to Elizabeth M. Brannon or Jessica F. Cantlon, Center for Cognitive Neuroscience, Duke University, Box 90999, Durham, NC 27708–0999. E-mail: brannon@duke.edu, cantlon@duke.edu

caterpillars and two tennis rackets look nothing alike and yet both exemplify a set of two. Similarly, if asked to count the number of animals in a room, three cats would be no better an example of threeness than a mouse, a cat, and a dog. It is this second sense of numerical abstraction on which we focus in this article.

Like human adults, human infants and nonhuman animals also represent number although they obviously do so in the absence of language (for reviews see Brannon, 2005; Dehaene, 1997; Feigenson, Dehaene, & Spelke, 2004; Gallistel & Gelman, 1992; Wynn, 1995). There are at least two different nonverbal systems used by infants and animals to represent number. One system functions to represent number as a continuous magnitude and discrimination in this system obeys Weber's law (analog magnitude number system; see Brannon & Terrace, 2002; Hauser, Tsao, Garcia, & Spelke, 2003; Lipton & Spelke, 2003; Platt & Johnson, 1971; Xu & Spelke, 2000). Evidence is accumulating that infants and perhaps nonhuman animals also possess a second system for representing small numerosities nonverbally as independent object files (object file system; see Feigenson & Carey, 2003; Hauser, Carey, & Hauser, 2000; Uller, Carey, Huntley-Fenner, & Klatt, 1999).

An interesting question is whether any given mechanism of number representation can be applied broadly to any and all objects or events. Gelman and Gallistel (1978) proposed that a key attribute of counting is the abstraction principle, which states that one can count any discrete entity regardless of its identity (apples, people, buildings, sounds, etc.). Of course it is conceivable that some organisms might be able to represent number but that they only can do so for some restrictive set of ecologically relevant objects. It is, however, clear that adult humans are just as adept at counting a random array of objects as they are at enumerating a set of red apples. It is therefore important to determine when this ability emerges in development, ontogenetically and phylogenetically.

Gelman and Gallistel (1978) also argued that heterogeneity of elements in a to-be-counted array might actually facilitate counting by allowing one to keep track of which elements had been counted and which were left to count (e.g., red ones done, now enumerate blue ones). Heterogeneity might also aid object individuation, which is a first step to any enumeration process. For example, individuation of adjacent objects in infancy is aided by shape, texture, and color distinctions between the objects (e.g., Needham, 1999). Thus heterogeneity might actually benefit any nonverbal iterative enumeration process where attention is used to select the items to be enumerated serially. Alternatively, heterogeneity might lead to overestimates of number or otherwise impede the enumeration process if heterogeneity led to an illusion of greater density. For example, Barth, Kanwisher, and Spelke (2003) suggested that adult humans form nonverbal numerical representations through a noniterative process whereby an estimate of cumulative area is multiplied by an estimate of density (see also Church & Broadbent, 1990). If heterogeneity of element size and shape interfered with an organism's ability to calculate cumulative area or density, heterogeneity might then lower the reliability of such a noniterative process as an estimate of number. This could be especially true for human infants who appear to have difficulty representing the cumulative area of large arrays of discrete elements (e.g., Brannon, Abbott, & Lutz, 2004).

To assess whether a young child's number representations obey the abstraction principle, a handful of researchers have investigated whether children can numerically match arrays when they are composed of heterogeneous rather than homogeneous elements. In one of the earliest studies of this sort, Von Gast (1957) found that preschoolers accurately labeled the cardinality of arrays of one to five homogeneous items, but failed to do so with heterogeneous items. Similarly, Siegel (1974) found that 4- and 5-year-old children took significantly longer to reach a performance criterion when required to match the numerals 1, 2, or 3 to the corresponding arrays if elements within an array were heterogeneous, rather than homogeneous in shape and color. By contrast, Beckmann (1924) found that 3- and 4-year-olds were able to label the cardinality of arrays of one to four heterogeneous items. Similarly, Gelman and Tucker (1975) found that heterogeneity had no effect on estimator or operator processes. In their first experiment, they found that children were equally proficient at providing cardinal labels for heterogeneous and homogeneous arrays of one to five elements. In their second experiment they demonstrated that children's assessments of whether number had changed or remained constant were largely uninfluenced by color or identity changes of a subset of the elements in a set.

Most recently, Mix (1999) found that the ability to nonverbally match stimuli based on numerical equivalence emerged gradually between 3 and 4 years of age and was aided by increasing physical similarity between arrays. Specifically children were better at matching two-, three-, and four-dot arrays to arrays of plastic discs than to arrays of shells or heterogeneous collections of objects. She argued that this developmental pattern suggests that young children require superficial similarities between arrays to notice abstract and conceptual commonalities such as number and that language affords the developing child an abstract number representation that is not tied to the nonnumerical perceptual attributes of stimuli.

Heterogeneous arrays have also been used in various investigations of animal numerical ability (e.g., Emmerton, 1998; Koehler, 1951; Murofushi, 1997; Nieder, Freedman, & Miller, 2002). However the main goal of these previous studies was to increase stimulus variability and thereby reduce the likelihood that animals would use nonnumerical features, such as surface area or perimeter, rather than to investigate the abstraction principle. Thus, within-stimulus variability was therefore not varied systematically in these studies. The general finding of these previous experiments was that accuracy was slightly lower for heterogeneous arrays as compared with homogeneous arrays.

Using a different paradigm, Brannon and Terrace (1998, 2000) trained two rhesus monkeys to respond to both homogeneous and heterogeneous visual arrays in ascending numerical order. This experiment has since been replicated with a

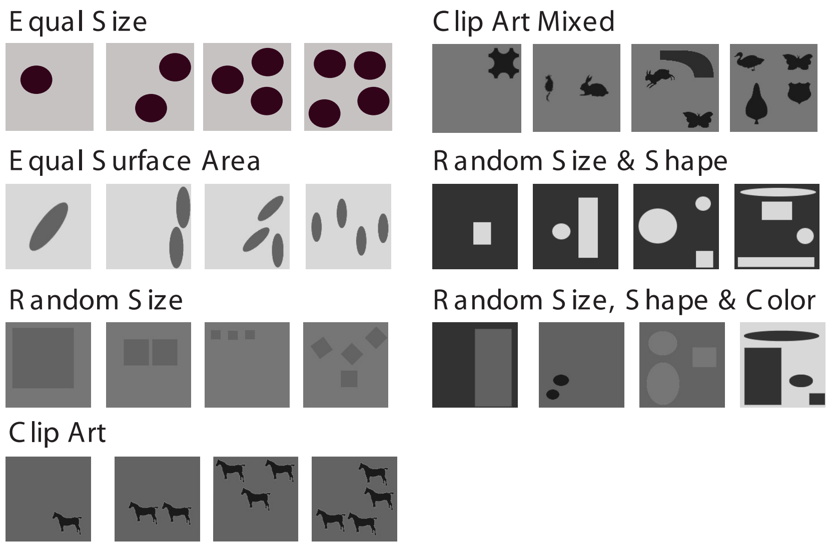

FIGURE 1 Exemplars of the seven different types of stimulus sets used by Brannon and Terrace (1998). Equal size: Elements were of same size and shape. Equal area: Cumulative area of elements was equal. Random size: Element size varied randomly across stimuli. Clip art: Identical nongeometric elements selected from clip art software. Clip art mixed: Clip art elements of variable shape. Random size and shape: Elements within a stimulus were varied randomly in size and shape. Random size, shape, and color: Same as previous with background and foreground colors varied between stimuli. From Brannon and Terrace (1998). Reprinted with permission from *Science, 282*, 746–749. Copyright 1998 AAAS.

baboon, a squirrel monkey, and three cebus monkeys (Judge, Evans, & Vyas, 2005; Smith, Piel, & Candland, 2003). The monkeys were required to touch, in ascending numerical order, the stimulus with one, two, three, and four elements. All the monkeys ordered trial-unique novel sets with above-chance accuracy, demonstrating their ability to represent the numerosities 1 through 4. However, performance was not equivalent for the seven stimulus types (see Figure 1).

Table 1 shows accuracy on the 150 trial-unique sets for each of the seven stimulus classes for the 7 individuals of the four monkey species that have been tested in this paradigm (Brannon & Terrace, 2000; Judge et al., 2005; Smith et al., 2003). Figure 2 shows the average accuracy for all monkeys as a function of stimulus category. On average, the monkeys performed better on the four homogeneous stimulus classes (39.6%) compared with the three heterogeneous stimulus classes (29%), $t(6) = 2.8$, $p < .05$. Indeed, six of the seven monkeys performed better on homogeneous compared with heterogeneous stimulus types. Accordingly, Brannon and Terrace's (2000) study suggested that numerical ordering is more difficult with heterogeneous arrays than homogeneous arrays for monkeys. There were, however, some exceptions to this generalization. For example,

TABLE 1
Percentage Correct as a Function of Stimulus Type for Four Monkey Species

Stimulus Class	*Rhesus 1*	*Rhesus 2*	*Baboon*	*Saimiri*	*Cebus 1*	*Cebus 2*	*Cebus 3*
Equal size	60 (1)	57 (2)	24 (4)	43 (1)	67 (1)	62 (1)	24 (4.5)
Equal surface area	56 (2)	59 (1)	50 (1)	23 (2)	64 (2)	45 (3.5)	45 (1)
Random size	50 (3)	23 (6)	45 (2)	9 (5)	41 (4)	41 (4)	14 (6)
Heterogeneous random size	22 (7)	38 (5)	21 (5)	10 (4.5)	42 (3)	45 (3.5)	24 (4.5)
Homogeneous clip art	31 (5)	50 (3)	27 (3)	14 (3)	36 (6)	23 (5)	27 (3)
Heterogeneous clip art	47 (4)	48 (4)	19 (6)	10 (4.5)	38 (5)	48 (2)	29 (2)
Heterogeneous random size and color	28 (6)	19 (7)	4 (7)	5 (6)	na	na	na

Note. Data taken from Brannon & Terrace (2000), Smith et al. (2003), and Judge, Evans, & Vyas (2005). Numbers in parentheses refer to standard errors across sessions.

there was no clear advantage in the homogeneous clip art condition compared with the heterogeneous clip art condition, $t(6) = .99$, $p > .05$. Four of the seven monkeys showed little difference in performance between these two classes, scoring within 4% of each other and two of the three remaining monkeys actually performed better in the heterogeneous compared with the homogeneous clip art condition. Similarly, there was no difference in performance on the homogeneous random size condition compared with the heterogeneous random size condition, $t(6) = .48$, $p > .05$.

The purpose of this experiment was to systematically address the question of whether rhesus monkeys have more difficulty evaluating numerical ordinal relationships between visual arrays composed of heterogeneous compared with homogeneous elements. Our study differs from that of Brannon and Terrace (1998, 2000) and subsequent replications (Judge et al., 2005; Smith et al., 2003) in that monkeys in this study were trained exclusively on stimuli that contained homogeneous elements. Subsequently, in testing, within-stimulus heterogeneity was systematically manipulated by varying the color, size, and shape of the elements. If within-stimulus heterogeneity affects a monkey's ability to form numerical representations, a graded level of accuracy should arise whereby performance decreases as the number of stimulus dimensions varied increases.

METHOD

Subjects and Apparatus

Subjects were two adult rhesus macaques (Mikulski and Feinstein), both of whom were female. Monkeys were socially housed with two other female rhesus

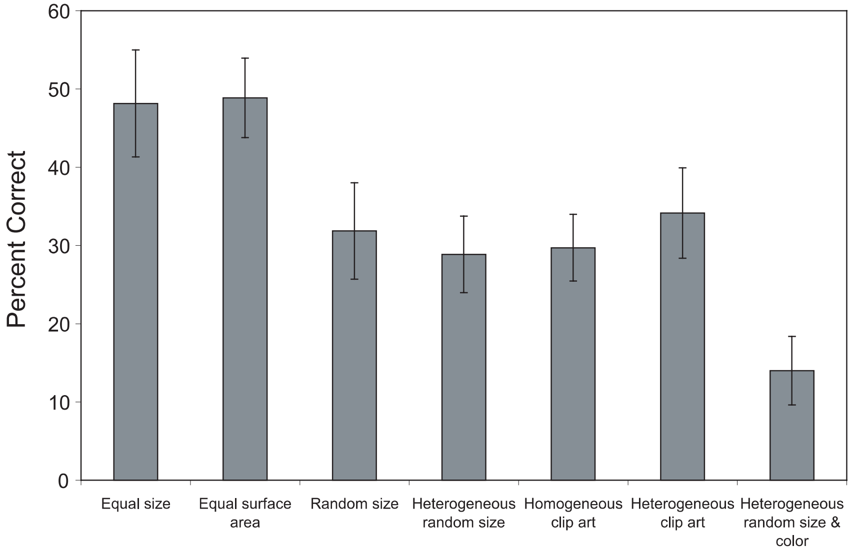

FIGURE 2 Percentage correct for seven monkeys on each of seven stimulus classes. Error bars reflect standard errors. Data from Brannon and Terrace (1998); Smith et al. (2003); Judge et al., 2005.

macaques. Monkeys were kept on controlled access to water to increase motivation for juice reward. All animal procedures followed an IACUC (Institutional Animal Care and Use Committee) protocol.

Monkeys were trained to be easily removed from their cages and sit in primate chairs in a soundproof booth that was located in a room adjacent to the colony room. Stimuli were presented on a touch-sensitive screen and were rewarded with small squirts of fruit juice. A Real Basic program was used to present stimuli and record responses.

Both rhesus monkeys had approximately 6 months of prior experience on the task used in this experiment (Cantlon & Brannon, in press, 2005). In these training sessions, animals were required to touch, in an ascending order, each member of a pair of numerosities with a value ranging from 1 to 9.

Stimuli

Eight stimulus conditions were used. Examples of each condition are shown in Figure 3. In the homogeneous condition, all elements within a stimulus were identical in size, shape, and color. In the heterogeneous color condition, elements were identical in size and shape, but varied in color. In the heterogeneous size condition,

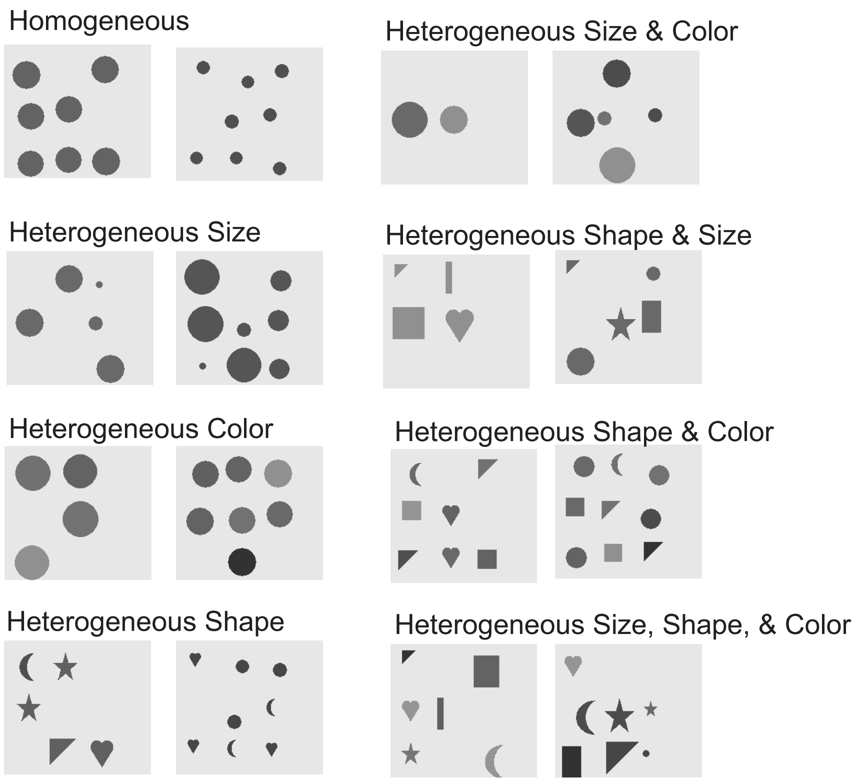

FIGURE 3 Example stimulus sets. Homogeneous: All elements within a stimulus were identical in size, shape, and color. Heterogeneous color: Elements were identical in size and shape, but varied in color. Heterogeneous size: Elements were identical in color and shape, but varied in size. Heterogeneous shape: Elements were identical in size and color but varied in shape. Heterogeneous color and size: Elements were identical in shape, but varied in both size and color. Heterogeneous color and shape: Elements were identical in size but varied in color and shape. Heterogeneous size and shape: Elements were identical in color, but varied in size and shape. Heterogeneous size, color, and shape: Elements varied in size, color, and shape.

elements were identical in color and shape, but varied in size. In the heterogeneous shape condition, elements were identical in size and color but varied in shape. In the heterogeneous color and size condition, elements were identical in shape, but varied in both size and color. In the heterogeneous color and shape condition, elements were identical in size but varied in color and shape. In the heterogeneous size and shape condition, elements were identical in color, but varied in size and shape. Finally, in the heterogeneous size, color, and shape condition, elements varied in size, color, and shape. Only within-stimulus heterogeneity was varied; in all conditions both stimuli were from the same condition (i.e., both varied on 0, 1, 2, or 3 dimensions).

Elements could be one of 15 different colors, five different sizes, and six different shapes (circles, stars, moons, rectangles, triangles, and hearts). The cumulative surface area of the elements ranged from 100 to 2,000 pixels. When a given dimension was varied, each element was assigned randomly to one of the possible values with replacement. In contrast, if the stimulus was homogeneous for a given dimension, a single value was chosen for all elements from the possible values, except for stimuli that were homogeneous in shape. These were always circles. Elements within each stimulus were randomly placed on a 180 × 180 pixel yellow background and each stimulus appeared in one of nine locations (3 × 3 grid).

Surface area controls. The smaller numerosity had a smaller cumulative surface area on 50% of all trials. When the smaller numerosity had the larger cumulative area it was on average 4,204 pixels and the larger numerosity was on average 2,212 pixels. When the larger numerosity had the greater area it was on average 6,581 pixels and the smaller numerosity was on average 2,416 pixels. Cumulative surface area ranged from 100 to 20,000 pixels for the smaller numerosity and from 200 to 22,500 pixels for the larger numerosity. The cumulative perimeter of the elements in a given stimulus was greater for the smaller numerosity on 25% of the trials. Cumulative perimeter ranged from 40 to 1,600 pixels for the smaller numerosity and from 80 to 1,800 pixels for the larger numerosity.

Task and Protocol

The task was to first touch the stimulus that contained the smaller number of elements and then touch the other stimulus. Each trial was initiated by touching a start stimulus (black square on a red background) in the bottom right corner of the screen. Reaction time to the first stimulus was the time it took the monkey to press the smaller numerosity after the start stimulus disappeared. The intertrial interval was fixed at 1 sec. Visual and auditory feedback followed each correct (border flash around stimulus and chime sound) or incorrect response (black screen flash and warning sound). Correct sequences were rewarded with .3 mL of juice. Incorrect responses to the first stimulus caused the trial to end and initiated a 3-sec timeout (TO). The chance probability of a correct response in this task was .5.

The first block of sessions contained four of the eight stimulus conditions: homogeneous, heterogeneous color, heterogeneous size, and heterogeneous color and size (shape remained homogeneous throughout all conditions of the first block). The second block of sessions included all eight conditions. To allow a fair comparison among the eight conditions it was necessary to equate the representation of each of the 36 numerosity pairs in each of the eight conditions. Only the first five trials for each of the 36 numerosity pairs for each of the eight conditions were included in the accuracy by condition analyses because this was the minimum number of trials each monkey completed on a given number pair (1,440 trials per monkey).[1] It was necessary to equate the representation of each numerosity pair for two reasons. First, to assess accuracy across conditions it was necessary to equate the number of trials per numerosity pair so that particularly easy or difficult pairs could not have been overrepresented and artificially inflated or deflated accuracy in that condition. Second, to analyze accuracy as a function of time in the experiment it was necessary to ensure that there was equal representation of each numerosity pair across successive blocks of trials.

The first three correct responses for each of the 36 numerosity pairs for each of the eight conditions were analyzed for latency to the first response. This number of trials was chosen because this was the minimum number of correct responses available for every numerosity pair. The reasons for equating the contribution of each numerosity pair in each condition and each trial block in the latency analyses were the same as described earlier for the accuracy analyses.

RESULTS

Overall, both monkeys performed very well in the ordinal comparison task, completing an average of 79% of the trials correctly (82% for Feinstein and 76% for Mikulski). Heterogeneity had an almost negligible effect on accuracy (as shown in Table 2). An analysis of variance (ANOVA) on monkeys' accuracy with a between-subject factor of monkey (Mikulski and Feinstein) and within-subject factor of experimental condition (eight levels) revealed only a main effect of monkey, $F(1, 8) = 16.15$, $p < .05$, and no main effect of condition or interaction. The main effect of subject was due to higher performance across all conditions for Feinstein.

Furthermore, accuracy did not improve across the five blocks of trials. An ANOVA with factors of subject and block revealed only a main effect of subject, $F(1, 14) = 11.86$, $p < .05$. This suggests that subjects were not learning to order heterogeneous sets within the course of the experiment but instead that there was immediate transfer from training with homogeneous stimuli to testing with both homogeneous and heterogeneous stimuli. Further evidence in favor of this conclusion is that there was no advantage for homogeneous over heterogeneous

[1]The actual number of trials each monkey completed for each number pair ranged between 5 and 11.

TABLE 2
Percentage Correct as a Function of Stimulus Type

Stimulus Class	*Mikulski*	*Feinstein*
Homogeneous	80 (2)	87 (2)
Heterogeneous color	74 (6)	81 (5)
Heterogeneous size	77 (4)	88 (1)
Heterogeneous shape	81 (1)	78 (7.5)
Heterogeneous size and color	71 (7.5)	86 (3)
Heterogeneous size and shape	71 (7.5)	82 (4)
Heterogeneous shape and color	78 (3)	80 (6)
Heterogeneous size, shape, and color	76 (5)	78 (7.5)

Note. Numbers in parentheses refer to standard errors across sessions.

stimuli in any of the five blocks. A Condition × Block ANOVA revealed no effect of condition, $F(7, 600) = 1.60$, $p > .05$, or block, $F(4, 600) = 0.74$, $p > .05$, and no interaction, $F(28, 600) = 1.14$, $p > .05$.

Two additional analyses were conducted to determine whether heterogeneity posed any specific costs on accuracy. First, accuracy for each condition was averaged as a function of whether zero, one, two, or all three stimulus dimensions varied. For example, the heterogeneous color and heterogeneous size and heterogeneous shape conditions are all classified as one stimulus dimension varied. There was a slight tendency for accuracy to decrease with the number of dimensions varied, but this was not significant (see Figure 4). An ANOVA on accuracy with factors of subject and number of dimensions varied revealed only a main effect of subject, $F(1, 8) = 16.51$, $p < .01$.

Second, to determine whether heterogeneity within any of the three stimulus dimensions posed a particular difficulty for the monkeys, the conditions were averaged based on whether color, size, or shape was varied or constant. For example, all four conditions used in the first block contained stimuli that were homogeneous with respect to shape and so all would be considered shape constant despite the fact the color or size of elements in these conditions sometimes varied. Figure 5 shows that accuracy was unimpaired by color, size, or shape heterogeneity: $t(4) = .09$, $p > .017$[2]; $t(4) = 2.6$, $p > .017$; $t(4) = 1.13$, $p > .017$.

Latency to respond. An ANOVA on latency to respond with factors of subject and number of dimensions varied revealed only a main effect of subject, $F(1, 4) = 23.3$, $p < .01$ (Figure 4). The main effect of subject was caused by faster responses by Feinstein. An ANOVA on reaction time (RT) with a between-subject factor of monkey (Mikulski and Feinstein) and within-subject factor of experimental

[2]The Bonferroni corrected alpha is .05/3 = .017.

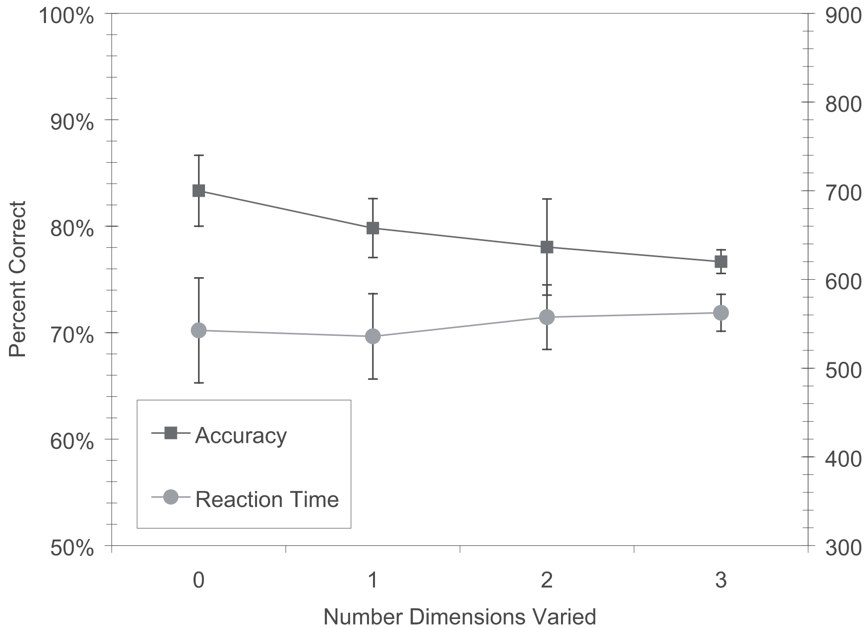

FIGURE 4 Accuracy (gray circles) and latency (black squares) as a function of the number of stimulus dimensions varied for and monkeys. (0) refers to the homogeneous condition; (1) refers to heterogeneous color, size, or shape conditions; (2) refers to heterogeneous size and color, heterogeneous shape and color, and heterogeneous size and shape; (3) refers to heterogeneous size, color, and shape. Error bars reflect standard errors of the mean (SEM).

condition (eight levels) revealed main effects of subject, $F(1, 4) = 37.87$, $p < .05$, and condition, $F(7, 28) = 3.85$, $p < .05$, and an interaction between subject and condition, $F(7, 28) = 2.89$, $p < .05$. The interaction was due to a lack of difference in RT between conditions for Mikulski in contrast to a difference in RT as a function of condition for Feinstein. Specifically, Feinstein was slightly faster when shape was homogeneous compared to when it was heterogeneous (471 msec vs. 540 msec), $t(3) = 6.76$, $p < .05$.

Accuracy as a function of stimulus controls. Table 3 shows that both monkeys' choices were significantly better than chance in all stimulus control conditions, demonstrating that they could accurately choose the smaller numerosity independently of cumulative surface area or cumulative perimeter. However, Feinstein showed a somewhat surprising pattern whereby her performance was significantly better on pairs for which cumulative surface area was incongruent

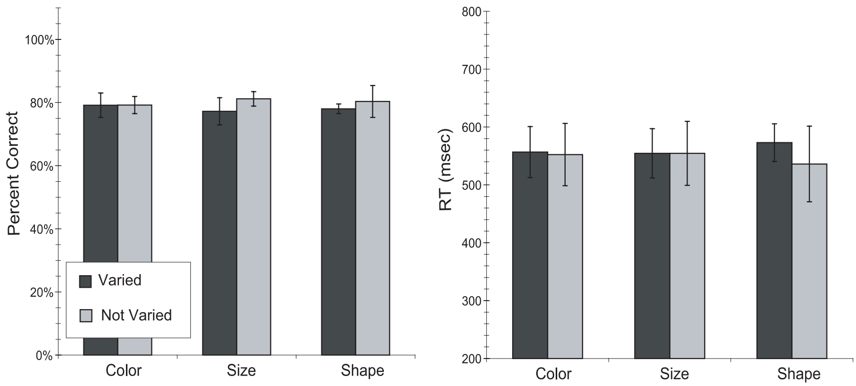

FIGURE 5 Top: Accuracy (left) and latency (right) on trials where color, size, or shape of elements within a stimulus were heterogeneous or homogeneous. Error bars reflect SEM.

with numerosity (e.g., smaller numerosity had larger surface area), $t(4) = 8.3$, $p < .025$[3]). Similarly, the same monkey was significantly more accurate when cumulative perimeter was incongruent with numerosity, $t(4) = 3.65$, $p < .025$. The second monkey showed no difference based on the cumulative area or cumulative perimeter of the stimuli: area, $t(4) = 1.00$, $p > .025$; perimeter, $t(4) = 2.83$, $p > .025$. The most important point for our current purposes is that accuracy far exceeded chance expectations regardless of the relation between number, surface area, and perimeter.

Distance effects. As in previous research the monkeys' accuracy and latency were modulated by the numerical distance between the stimuli within a pair (e.g., Brannon & Terrace, 1998, 2000; Cantlon & Brannon, in press, 2005). Figure 6 illustrates that accuracy increased with distance and latency to the first response decreased with distance. Figure 6 also shows that these distance effects did not vary as a function of stimulus heterogeneity. An ANOVA on accuracy with factors of subject, condition, and numerical distance revealed main effects of subject, $F(1, 7) = 15.08$, $p < .01$, and numerical distance, $F(7, 448) = 38.72$, $p < .01$, and no interactions. A parallel ANOVA on latency similarly revealed main effects of subject, $F(1, 7) = 48.93$, $p < .01$, and numerical distance, $F(7, 448) = 14.72$, $p < .01$, and no interactions.

[3]The Bonferroni corrected alpha for conducting two contrasts per subject is .05/2 = .025.

TABLE 3
Percentage Correct as a Function of Stimulus Control

	Feinstein	*Mikulski*
Surface area congruent	77	76
Surface area incongruent	90	75
Perimeter congruent	82	77
Perimeter incongruent	87	72

DISCUSSION

Our results suggest that the degree to which elements within a stimulus vary does not impose a cost or benefit when monkeys make ordinal numerical comparisons. A host of different analyses failed to find any robust effects of stimulus heterogeneity on monkeys' accuracy or latency to respond. In addition both monkeys' accuracy and latency to respond were modulated by the numerical disparity between the two numerosities compared, and this effect of distance held across all eight stimulus conditions regardless of the number of stimulus dimensions that varied.

The lack of a major cost of stimulus heterogeneity is particularly striking in light of the fact that both monkeys had extensive prior training with ordering numerical stimuli with homogeneous elements and had never been exposed to stimuli with heterogeneous elements. Despite this previous training with exclusively homogeneous arrays, which might have biased them to expect homogeneity, they showed no impairment in accuracy on stimuli that contained heterogeneous elements compared to stimuli that contained homogeneous elements. Furthermore, there was no evidence of improvement in accuracy on heterogeneous stimuli over the five experimental blocks, suggesting that monkeys did not have to learn to ignore variability in elements within a stimulus but instead that they did so immediately.

These findings demonstrate that at least one type of nonverbal number representation follows the abstraction principle. That is, the color, shape, or size of elements in an array seems to have no impact on the process by which monkeys form numerical representations. The mechanism registers an element regardless of its identity, allowing it to form an abstract representation of number. This is a significant finding, because although it was known that animals could form numerical representations for heterogeneous arrays, this is the first systematic investigation that shows that there is no cost to a shift from a homogeneous to a heterogeneous array.

These data provide further argument against Mix's (1999) hypothesis that mastery of the verbal counting system is necessary for one to ignore irrelevant surface features and attend to numerical equivalence. Although these data demonstrate that

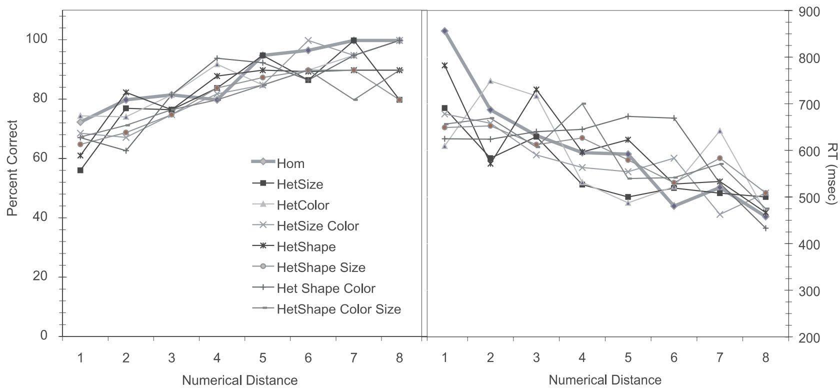

FIGURE 6 Average accuracy (left) and latency (right) as a function of numerical disparity for each of the eight conditions. The bolded function is the homogeneous condition.

nonverbal number representations provide a mechanism for representing number independently of element shape, color, or size, it is certainly possible that language plays an important role in allowing children to identify number as a relevant stimulus attribute.

Our data cannot address what role featural information plays in establishing numerical representation. Current models of nonverbal number representation propose that representations of number are largely devoid of featural information. For example, Leslie and colleagues have hypothesized that color and shape function to aid individuation but that identity information is not preserved in object file representations (Leslie, Xu, Tremoulet, & Scholl, 1998). Similarly, both the Meck and Church (1983) mode-control model and the Dehaene and Changeux (1993) neural network model, which result in analog magnitude representations of number, propose that the enumeration process functions independently of perceptual features of objects or events such as color, duration, or size. However, this leaves open the question of whether element features are not encoded at all when monkeys enumerate or whether they are encoded but ignored by the enumeration process. Future experiments that test whether monkeys retain element identity information while still forming numerical representations over heterogeneous elements would be necessary to distinguish between these alternatives.

More generally, the degree to which a monkey's numerical knowledge is conceptual as opposed to perceptual is not at all clear (Mandler, 2004). In Mandler's (2004) terms a concept refers to "declarative knowledge about object kinds and events that is potentially accessible to conscious thought" (p. 4). In contrast Gallistel and Gelman (1992) contended that an animal may be said to have a numerical concept "insofar as they may be shown to mentally manipulate numerons (nonverbal number representations) in processes that are isomorphic to some or all of the operations that define the system of arithmetic: ordering, addition, subtraction, multiplication, and division" (p. 45). Clearly, our data satisfy Gallistel and Gelman's definition of conceptual numerical knowledge, but less clear is whether a monkey's numerical knowledge satisfies Mandler's definition.

Although heterogeneity seems to have no impact on analog number representations, it is possible that heterogeneity differentially effects number representations built within the object file system. Recent data by Feigenson (2005) suggest that young infants attend preferentially to area over number when presented with small numerosities composed of homogeneous elements, but to number over area when elements are heterogeneous. Thus, heterogeneity may actually enhance an infant's ability to attend to the numerosity of small arrays. It will be of great interest to determine the effect of heterogeneity on the representation of both large and small numerosities in infants. Particularly interesting would be if heterogeneity had different effects on the object file and analog magnitude nonverbal number systems in infancy.

Many questions remain about the nature of nonverbal numerical representations and the impact of within-stimulus heterogeneity on mechanisms of forming nonverbal numerical representations. For example, how does stimulus heterogeneity affect numerical discrimination when elements are presented successively or in other modalities? Can monkeys appreciate the numerical correspondence between homogeneous or heterogeneous sets experienced in different modalities? Perhaps most profoundly, what is the relation between an infant or young child's ability to represent number nonverbally and the acquisition of the verbal counting system (e.g., Carey, 2004; Gelman & Meck, 1983). Answers to these questions will contribute toward a better understanding of the fundamental bases of nonverbal numerical thinking.

ACKNOWLEDGMENTS

This research was funded by RO1 HD049912-01A2, an NSF Career award, and a John Merck fellowship to Elizabeth M. Brannon and an NSF graduate fellowship to Jessica F. Cantlon. We gratefully acknowledge the help of Kerry Jordan, Kerrie Lewis, and Pierre Rojas in data collection with monkeys. We further thank Lisa Feigenson, Michael Platt, Jamie Roitman, Herb Terrace, and all members of the Brannon Lab for comments on a previous draft of this article.

REFERENCES

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

Beckmann, H. (1924). Die Entwicklung der Zahlleistung bei 2–6 jahrigen Kindern [The development of numerical achievement in 2–6 year-old children]. *Zietschrift fur Angewandte Psychologie, 22,* 1–72.

Brannon, E. M. (2005). What animals know about number. In J. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 85–108). New York: Psychology Press.

Brannon, E. M., Abbott, S., & Lutz, D. (2004). Number bias for the discrimination of large visual sets in infancy. *Cognition, 93,* B59–B68.

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

Brannon, E. M., & Terrace, H. S. (2000). Representation of the numerosities 1–9 by rhesus monkeys (*Macaca mulatta*). *Journal of Experimental Psychology: Animal Behavior Processes, 26,* 31–49.

Brannon, E. M., & Terrace, H. S. (2002). The evolution and ontogeny of ordinal numerical ability. In C. Allen, M. Bekoff & G. M. Burghardt (Eds.), *The cognitive animal* (pp. 197–204). Cambridge, MA: MIT Press.

Cantlon, J., & Brannon, E. M. (2005). Semantic congruity facilitates number judgments in monkeys. *Proceedings of the National Academy of Sciences, 102,* 16507–16511.

Cantlon, J., & Brannon, E. M. (in press). Shared system for ordering small and large numbers in monkeys and humans. *Psychological Science*.

Carey, S. (2004). Bootstrapping and the origin of concepts. *Daedalus, Winter*, 59–68.

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

Dehaene, S. (1997). *The number sense: How the mind creates mathematics*. 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.

Emmerton, J. (1998). Numerosity differences and effects of stimulus density on pigeons' discrimination performance. *Animal Learning and Behavior, 26,* 243–256.

Feigenson, L. (2005). A double-dissociation in infants' representations of object arrays. *Cognition, 95,* B37–B48.

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

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences, 8,* 307–314.

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

Gelman, R., & Gallistel, C. R. (1978). *The child's concept of number*. Cambridge, MA: Harvard University Press.

Gelman, R., & Meck, E. (1983). Preschooler's counting: Principles before skill. *Cognition, 13,* 343–359.

Gelman, R., & Tucker, M. F. (1975). Further investigations of the young child's conception of number. *Child Development, 46,* 167–175.

Hauser, M. D., Carey, S., & Hauser, L. B. (2000). Spontaneous number representation in semi-free-ranging rhesus monkeys. *Proceedings of the Royal Society, London, 267,* 829–833.

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

Judge, P. G., Evans, T. A., & Vyas, D. K. (2005). Ordinal representation of numeric quantities by brown capuchin monkeys (*Cebus apella*). *Journal of Experimental Psychology: Animal Behavior Processes, 31,* 79–94.

Koehler, O. (1951). The ability of birds to "count." *Bulletin of Animal Behavior, 9,* 41–45.

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

Mandler, J. M. (2004). *The foundations of mind.* New York: Oxford University Press.

Meck, W. H., & Church, R. M. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behavior Processes, 9,* 320–334.

Mix, K. S. (1999). Similarity and numerical equivalence: Appearances count. *Cognitive Development, 14,* 269–297.

Murofushi, K. (1997). Numerical matching behavior by a chimpanzee (*Pan troglodytes*): Subitizing and analogue magnitude estimation. *Japanese Psychological Research, 39,* 140–153.

Needham, A. (1999). The role of shape in 4-month-old infants' object segregation. *Infant Behavior and Development, 22,* 161–178.

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.

Platt, J. R., & Johnson, D. M. (1971). Localization of position within a homogeneous behavior chain: Effects of error contingencies. *Learning and Motivation, 2,* 386–414.

Siegel, L. S. (1974). Development of number concepts: Ordering and correspondence operations and the role of length cues. *Developmental Psychology, 10,* 907–912.

Smith, B. R., Piel, A. K., & Candland, D. K. (2003). Numerity of a socially housed hamadryas baboon (*Papio hamadryas*) and a socially housed squirrel monkey (*Saimiri sciureus*). *Journal of Comparative Psychology, 117,* 217–225.

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

Von Gast, H. (1957). Der umgang mit Zahlen und zahlgebilden in der fruhen kindheit [Number constructs in early childhood]. *Z. Psychol, 161,* 1–90.

Wynn, K. (1995). Origins of numerical knowledge. *Mathematical Cognition, 1,* 35–60.

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