Semantic congruity affects numerical judgments similarly in monkeys and humans

Jessica F. Cantlon* and Elizabeth M. Brannon*

Department of Psychological and Brain Sciences and Center for Cognitive Neuroscience, Duke University, Box 90999, Durham, NC 27708-0999

Edited by Charles R. Gallistel, Rutgers, The State University of New Jersey, Piscataway, NJ, and approved September 24, 2005 (received for review July 29, 2005)

Monkeys (*Macaca mulatta*) were trained to order visual arrays based on their number of elements and to conditionally choose the array with the larger or smaller number of elements dependent on a color cue. When the screen background was red, monkeys were reinforced for choosing the smaller numerical value first. When the screen background was blue, monkeys were reinforced for choosing the larger numerical value first. Monkeys showed a semantic congruity effect analogous to that reported for human comparison judgments. Specifically, decision time was systematically influenced by the semantic congruity between the cue ("choose smaller" or "choose larger") and the magnitude of the choice stimuli (small or large numbers of dots). This finding demonstrates a semantic congruity effect in a nonlinguistic animal and provides strong evidence for an evolutionarily primitive magnitude-comparison algorithm common to humans and monkeys.

nonhuman primates | numerical cognition | analog magnitude | distance effect

Humans and nonhuman animals discriminate numbers in a way that obeys the psychophysical tenets of Weber's law (e.g., refs. 1–6; see ref. 7 for review). That is, animals and humans are faster and more accurate at comparing two numerical values as the ratio between them (min/max) decreases. For humans, the same pattern of ratio-dependent performance emerges regardless of whether the numerical values are presented as Arabic numerals, arrays of dots, or sequences of tones (e.g., refs. 1, 8–10). This response pattern is taken to indicate that humans and animals represent approximate numerical values as imprecise mental magnitudes (e.g., refs. 4 and 7). Thus, animals and humans are thought to represent approximate numerical values in fundamentally the same way. However, the specific process by which numbers are compared in monkeys and humans has not been specified. In this study, we investigate whether monkeys show a response signature of adult human comparison judgments: the semantic congruity effect.

When adult humans are asked "Which is smaller: an ant or a rat?", they are much quicker to respond than when asked "Which is larger: an ant or a rat?" (e.g., refs. 11 and 12). In contrast, when adult humans are asked to compare two large animals, such as a cow and an elephant, they are much quicker to respond when asked "Which is larger?" than "Which is smaller?" This effect is known as the semantic congruity effect and has been reported for adult humans when they compare stimuli along a variety of continua, including the distance between two cities (13), line length (14), brightness (15), the intelligence of animals (16), adjectives of ordinal quality (e.g., good, fair, poor, or excellent; ref. 17), surface area (18), and Arabic numerals (19, 20). In all cases, the semantic relationship between the direction of the choice objective and the perceived magnitude of the to-be-compared entities affects the rapidity of human decision-making.

For numerical comparisons, the semantic congruity effect occurs in adults when the magnitude of the choice stimuli (small or large numbers) conflicts with the ordinal direction of the choice objective ("choose the smaller number" or "choose the larger number"). In a classic demonstration of the semantic congruity effect on adult human numerical comparisons, Banks *et al.* (19) presented subjects with a pair of Arabic numerals ranging from 1 to 9. On some trials, subjects were instructed to choose the smaller number, whereas on other trials, they were instructed to choose the larger number. When the Arabic numerals were both small numbers (e.g., 2 and 3), subjects identified the smaller number faster than the larger number; when the Arabic numerals were both large numbers (e.g., 7 and 8), subjects identified the larger number faster than the smaller number. Thus, semantic congruity between the ordinal term in the verbal question (smaller or larger) and the subjective magnitude of the comparison numerals along the test continuum affects humans' decision-making time.

Previous research has demonstrated that semantic congruity influences the decision-making process during the computation of the comparison rather than during stimulus encoding (12, 21). Shaki and Algom (12) presented adults with picture–word compounds consisting of animal pictures with the names of animals printed within the pictures. Subjects were asked to "choose the larger" or "choose the smaller" of the two animal pictures or two animal names. On congruent trials, the animal name matched the picture of the animal. However, on incongruent trials, the animal name and the animal picture did not match and were incongruent in size. For example, a picture of an elephant would have the word "cat" printed on it. A lack of congruence between the semantic content of picture–word stimuli produced Stroop-like interference when subjects either read the words or named the pictures of the compound stimuli (22).

Following the logic of Sternberg (23), the authors reasoned that if semantic congruity uses resources involved in the stages of stimulus encoding, Stroop and semantic congruity effects should interact in reaction time (RT). If, on the other hand, Stroop and semantic congruity effects occur at different stages of processing, their affects should be additive. The results indicated that trials in which the picture–word stimuli were incongruent produced increases in subjects' RTs that were typical of the Stroop effect; however, these increases in RT were additive and did not interact with the semantic congruity effect. Stroop interference from the incongruent compound stimuli was attributed to the input stages of processing, whereas interference from the semantic congruity between the choice objective and the choice stimuli was attributed to the comparative stage of processing. Similar results have been obtained by Cech (21), who used perceptually degraded stimuli to show that impairments during stimulus encoding are additive with the effect of semantic congruity and do not interact. Again, the semantic congruity effect appeared to be a consequence of the comparison process rather than stimulus encoding.

PSYCHOLOGY

Conflict of interest statement: No conflicts declared.

This paper was submitted directly (Track II) to the PNAS office.

Abbreviation: RT, reaction time.

*To whom correspondence may be addressed. E-mail: jfc2@duke.edu or brannon@duke.edu.

In short, current research with human subjects suggests that the semantic congruity effect is a signature of the particular algorithm used during mental comparisons. Thus, the semantic congruity effect provides an opportunity for investigating the similarity between the comparative processes of monkeys and humans apart from the way they encode and represent stimuli. By investigating whether monkeys show a semantic congruity effect when they make numerical comparisons, we can determine whether monkeys not only represent numbers as analog magnitudes in much the same way as humans, but also whether they use the same comparison algorithm for deciding which numerical value is smaller or larger.

Although nonhuman animals and humans both possess a nonverbal system for representing number approximately (2, 3, 6, 24), only adult humans have the ability to represent exact numerical values as discrete mental quantities through linguistic symbols (25, 26). Because the semantic congruity effect has been reported only for human comparative judgments, it is possible that the effect is specific to the uniquely human ability to represent number precisely. Indeed, some models of human comparative processes have emphasized the role of symbolic and linguistic interference in producing the semantic congruity effect (e.g., refs. 11 and 16). These models assume that the semantic congruity effect arises when humans represent the elements of a given problem as discrete mental symbols that are analogous to words. In contrast, other models have emphasized the role of nonsymbolic, analog mental codes in comparative processes (14, 20). These models assume that the semantic congruity effect arises when the elements of a problem are compared along a continuous distribution rather than as discrete symbols. Evidence of a semantic congruity effect on monkeys' numerical comparisons would provide support for models emphasizing the role of analog magnitude representations over linguistic representations in producing the effect.

In addition to the effect of semantic congruity on performance, adults also are typically faster and more accurate overall at choosing the larger of two stimuli than at choosing the smaller of the same two stimuli (e.g., refs. 19 and 27). This main effect of ordinal direction has been attributed to linguistic biases in the use of comparative adjectives (linguistic markedness; refs. 27–30). In English, the default adjective for comparing particular magnitudes is "larger" rather than "smaller." That is, English speakers may be more likely to describe 10 as being larger than 5 rather than describing 5 as being smaller than 10. This asymmetry in the use of the terms "larger" and "smaller" has been proposed as the cause of humans' better performance when deciding which of two values is larger than which of two values is smaller. If a nonverbal animal also shows superior performance for choosing the larger of two stimuli, this finding would suggest that the effect is not dependent on language.

Here, we test whether monkeys, like humans, are influenced by semantic congruity when making approximate numerical judgments. If semantic congruity influences monkeys' numerical judgments as it influences humans' numerical judgments, the semantic congruity effect cannot be attributed to linguistic processes. Rather, evidence of a semantic congruity effect in animals that lack a discrete number system would implicate an analog magnitude-comparison algorithm shared by humans and nonhuman animals. Further, this result would be strong evidence that human computational processes can show evolutionary continuity with the computational processes of other primates.

Method

Subjects. Subjects were two adult female rhesus macaques (*Macaca mulatta*), Feinstein and Mikulski, who were socially housed along with two other rhesus macaque females. Both monkeys were kept on water-restricted diets that were approved by an institutional animal care and use committee to increase motivation for the juice reinforcement that was used in this study.

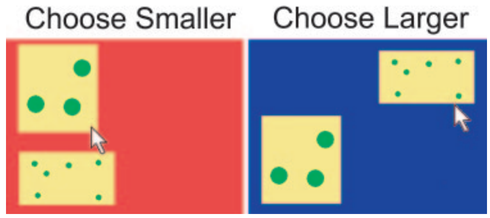

Fig. 1. Illustration of the task design. When the background color was red, monkeys were rewarded for pressing the stimulus with the smaller number of dots. When the background was blue, monkeys were rewarded for pressing the stimulus with the larger number of dots.

Task. Monkeys were tested in sound-attenuated booths while seated in Plexiglas primate chairs. Stimuli were presented on a 2×2 touch-screen matrix in randomly selected locations. To begin a trial, subjects were required to press a start stimulus, a small black square presented in the bottom left corner of the screen. After this response, two arrays of dots that each contained 1–9 elements were simultaneously presented. When the touch-screen background was red, the monkeys were rewarded for selecting the array with the smaller number of items ("choose smaller" trials), and when the background color was blue, the monkeys were rewarded for selecting the array with the larger number of items ("choose larger" trials) (see Fig. 1). The choose smaller and choose larger trials were randomly intermixed within a session.

The trial terminated after the subject made an incorrect response to the first item or after both items were responded to in the correct order. Correct sequences were followed by positive visual and auditory feedback and 0.3 ml of Kool-Aid (Kraft Foods, Northfield, IL) reinforcement. Incorrect responses were followed by a warning tone and the presentation of a black screen for 3–6 sec. There was a 0.5 probability of producing a correct response by chance. A movie of a monkey performing this task is available from the authors on request.

Training. Subjects were initially trained with a small subset of the numerosities 1–9. During the first 25 sessions of training, Feinstein was presented only with the numerosities 1, 3, 5, and 9, whereas Mikulski was trained only with the numerosities 1, 4, 5, and 9. The training subset was gradually expanded until it included all 36 possible numerical pairings of the numerosities between 1 and 9. During the first 20 sessions of training, choose smaller and choose larger trials were presented in alternating session blocks; however, after these initial 20 sessions, choose smaller and choose larger trials were randomly intermixed in each session with equal probability of occurrence. The training period consisted of $\approx$100 sessions, and each session consisted of $\approx$800 trials.

Testing. Monkeys were tested by using the general task and procedure described above. Monkeys were tested on $\approx$4,000 trials over five test sessions.

Stimuli. Training stimuli consisted of multiple exemplars of each of the numerosities 1–9. These stimuli were composed of elements of various shapes, sizes, and colors to encourage subjects to disregard these features when making their responses.

Stimuli presented during testing were novel exemplars of all 36 possible pairs of the numerical values between 1 and 9. There were 18 different exemplars of each numerical value, each of which was randomly combined with every other exemplar of the numerical values tested. All stimuli were tested in both choose smaller and choose larger ordinal directions.

To control for density, defined as the total number of elements per cm^2 of stimulus background space, density was sometimes congruent (≈62.5% of trials) and sometimes incongruent (≈37.5% of trials) with number. For example, when density was incongruent with number, density was greater in the stimulus with the smaller number of elements. The cumulative surface area of the elements and the cumulative perimeter of the elements in each array also were incongruent with number on 37.5% of trials.

Results and Discussion

Monkeys used the numerical values of the arrays as the basis for comparing stimuli on both the choose smaller and choose larger trials as opposed to surface area, density, or cumulative perimeter. Monkeys' accuracy and RT were modulated by the ratio between the numerical values compared, indicating that they represented number as analog magnitudes. Importantly, monkeys showed a semantic congruity effect in their decision time: they were faster to choose the smaller of two small numerical values but were faster to choose the larger of two large numerical values.

Overall, monkeys performed significantly above chance across the five testing sessions on both the choose smaller trials [single-sample t tests, Feinstein, mean accuracy 86%, $t(4) = 37.90, P < 0.01$; Mikulski, mean accuracy 71%, $t(4) = 21.45, P < 0.01$] and the choose larger trials [single-sample t tests: Feinstein, mean accuracy 92%, $t(4) = 52.31, P < 0.01$; Mikulski, mean accuracy 85%, $t(4) = 48.79, P < 0.01$]. Remarkably, both monkeys were able to conditionally select the smaller or larger of the two stimuli based on the color of the screen background. Achieving this level of performance is no small feat because it requires monkeys to make opposite responses to identical numerical pairs within a single session. For example, if subjects were presented with the pair 1-9 on a red background, they should select 1, but when presented with this pair on a blue background, they should select 9.

Although both monkeys were able to respond appropriately on choose larger and choose smaller trials, they both performed significantly better across numerical pairs presented during choose larger trials than during choose smaller trials [Mikulski, $t(35) = 7.41, P < 0.01$; Feinstein, $t(35) = 5.25, P < 0.01$]. Mikulski also was significantly faster at making choose larger than choose smaller comparisons [$t(35) = 2.49, P < 0.05$]; however, Feinstein responded equally fast in both ordinal directions [$t(35) = 0.33, P = 0.74$]. Increased speed and accuracy on choose larger trials over choose smaller trials also has been reported for adult humans tested on analogous numerical tasks (e.g., ref. 19). As reviewed above, performance asymmetries in human magnitude comparisons have been interpreted as a consequence of the frequencies with which people use comparative adjectives. For example, English speakers use the term "larger" more often than the term "smaller" to express relative size (e.g., refs. 27–30). Our finding of similar asymmetrical performance in monkeys suggests that better performance in the choose larger direction may not be caused by linguistic biases in humans.

Monkeys performed significantly above chance when the dimensions of density, cumulative surface area, and cumulative perimeter were congruent with number [single-sample t tests; Feinstein, $t(14) = 75.29, P < 0.001$; Mikulski, $t(14) = 89.37, P < 0.001$] and when they were incongruent with number [single-sample t tests; Feinstein, $t(14) = 71.55, P < 0.001$; Mikulski, $t(14) = 48.37, P < 0.001$]. Monkeys' performance on these three stimulus conditions is shown in Fig. 2. An ANOVA for Subject (Feinstein, Mikulski) × Stimulus Control (Density, Surface Area, Perimeter) × Condition (Congruent or Incongruent with Number) with session as the repeated measure revealed a main effect of Subject [Feinstein was significantly more accurate than Mikulski, $F(2, 24) = 641.7, P < 0.001$] and an interaction between Subject and Control Condition [$F(1, 24) = 5.8, P < 0.05$]. Fisher's least significant difference post-hoc tests revealed that Feinstein performed equally well ($P = 0.44$) on congruent (mean accuracy 89%) and incongruent trials (mean accuracy 90%), whereas Mikulski performed significantly better ($P < 0.05$) on congruent trials (mean accuracy 79%) than incongruent trials (mean accuracy 77%). Although one monkey did perform significantly (albeit by only 2%) better on congruent trials than incongruent trials, the more important point for our purposes is that accuracy for both monkeys was significantly above chance for incongruent trials, indicating that they were using the number of elements in the array as the basis for making comparisons rather than density, cumulative surface area, or cumulative perimeter.

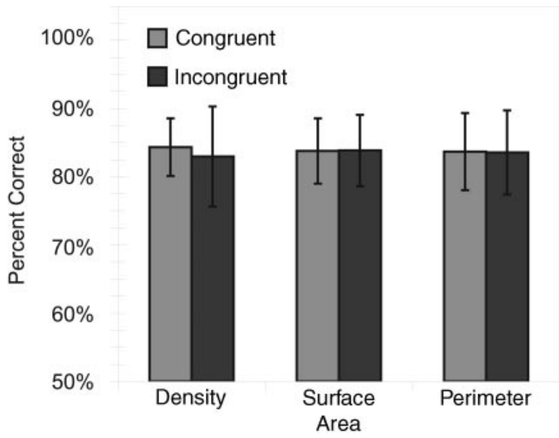

Fig. 2. Accuracy when the number of dots in a stimulus was congruent (smaller number has smaller density, perimeter, or surface area) and incongruent (larger number has smaller density, perimeter, or surface area or when stimuli are equated on dimension) with density, cumulative surface area, and cumulative perimeter. Accuracy was not affected by these stimulus controls. The error bars reflect standard error between monkeys.

As in many previous studies of numerical comparison judgments in animals, accuracy and RT were systematically influenced by the ratio (min/max) between the numerical values being compared [accuracy (Feinstein, $r^2 = -0.86, P < 0.05$; Mikulski, $r^2 = -0.74, P < 0.05$) and RT (Feinstein, $r^2 = 0.78, P < 0.05$; Mikulski, $r^2 = 0.79, P < 0.05$)]. The RT distance effects (Fig. 3*A*) held for each monkey for both the choose smaller

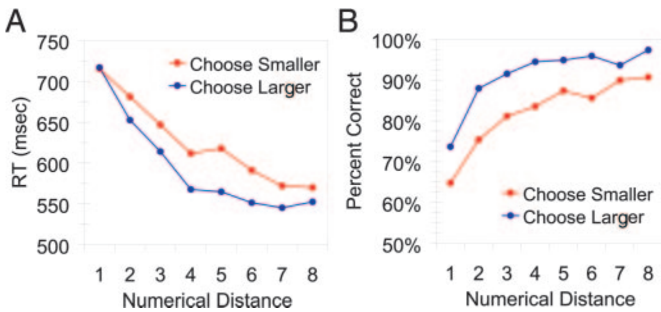

Fig. 3. The numerical distance effect. (*A*) RT as a function of numerical distance between the numerosities in a pair for all 36 numerical pairs during choose larger and choose smaller trials. Monkeys were faster as the distance between numerical values increased. (*B*) Accuracy as a function of numerical distance for all 36 numerical pairs during choose larger and choose smaller trials. Monkeys were more accurate as numerical distance increased.

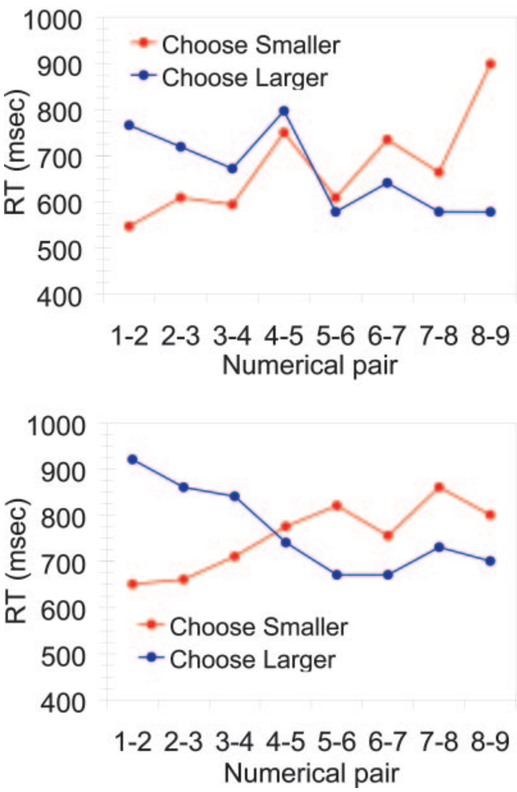

Fig. 4. The semantic congruity effect. RT for Feinstein (*Upper*) and Mikulski (*Lower*) for each numerical pair at a constant distance of 1 during choose smaller and choose larger trials. The crossover pattern reflects the effect of semantic congruity on monkeys' numerical comparisons.

(Feinstein, $r^2 = 0.84$, $P < 0.05$; Mikulski, $r^2 = 0.94$, $P < 0.05$) and choose larger trials (Feinstein, $r^2 = 0.84$, $P < 0.05$; Mikulski, $r^2 = 0.72$, $P < 0.05$). Similarly, accuracy on numerical comparisons increased as numerical distance increased (Fig. 3*B*) for both the choose smaller (Feinstein, $r^2 = 0.76$, $P < 0.05$; Mikulski, $r^2 = 0.88$, $P < 0.05$) and choose larger (Feinstein, $r^2 = 0.60$, $P < 0.05$; Mikulski, $r^2 = 0.63$, $P < 0.05$) trials. This aspect of our data confirms the well established finding that monkeys represent number as analog magnitudes (1, 3–6).

The main finding was that both monkeys showed a semantic congruity effect (Fig. 4).[†] Monkeys were 186 msec faster on choose smaller trials than on choose larger trials when the two values were small (e.g., 1 vs. 2). In contrast, monkeys were, on average, 136 msec faster on choose larger trials than on choose smaller trials when the two values were large (e.g., 8 vs. 9). An ANOVA for Subject (Feinstein, Mikulski) × Response Type ("choose smaller," choose larger) × Numerical Magnitude [Small (1 vs. 2, 2 vs. 3, 3 vs. 4), Large (6 vs. 7, 7 vs. 8, 8 vs. 9)][‡] on monkeys' RT to pairs of numerosities that differed by 1 revealed a main effect of Subject [$F(1, 8) = 20.32$, $P < 0.01$], reflecting overall faster RTs for Feinstein (mean, 667 msec) than Mikulski (mean, 764 msec) and an interaction between the Response Type and the Numerical Magnitude of the comparison stimuli [$F(1, 8) = 38.55$, $P < 0.001$]. The interaction reflects that, when the to-be-compared arrays were both small numbers, the monkeys were faster to choose the smaller than to choose the larger of two arrays [$t(5) = 5.60$, $P < 0.01$], whereas when both arrays contained a large number of dots, monkeys were faster to choose the larger array than to choose the smaller of two arrays [$t(5) = 3.61$, $P < 0.05$].

The response patterns of monkeys on this task are remarkably similar to those reported for human subjects by Banks *et al.* (19) using Arabic numeral stimuli in a paradigm similar to ours. In light of these previous studies with humans, our data suggest that the cognitive process that monkeys use to compare the numerical value of two arrays of dots shares important features with the process used by humans for Arabic numeral comparisons.

General Discussion

In summary, our results confirm that monkeys represent numerical values as analog magnitudes and demonstrate that semantic congruity affects the speed with which they compare numerical values. The qualitative similarity between the monkeys' response patterns and the response patterns of humans tested on analogous tasks with Arabic numerals (11, 19, 20) suggests that humans and monkeys are using the same mental-comparison algorithm for determining the larger or smaller of two numerical values.

Although language may have profound effects on some aspects of human thought (31), the fact that monkeys show a semantic congruity effect is evidence that certain computational processes have not been qualitatively altered by the emergence of language in humans. Although it is possible that the semantic congruity effect is a consequence of different psychological processes in monkeys and humans, this interpretation of the data seems less parsimonious than the notion that the effect results from a common, nonlinguistic comparative process in humans and monkeys. Thus, our study suggests that the computational process by which humans compare the relative magnitude of numerical values is evolutionarily conserved and language-independent.

Because monkeys do not represent numerical values as discrete quantities, the effect of semantic congruity on their response patterns cannot be explained by semantic conflicts that are hypothesized to result when humans mentally translate problems into linguistic symbols (e.g., ref. 19). There are several competing models of comparative processes that can account for the semantic congruity effect (see ref. 32 for review). Although our data cannot differentiate among these models beyond excluding those based on discrete representations, one model that could apply to both humans and nonhumans has been presented by Holyoak (20) and Petrusic (32). According to this hypothesis, the numerical values of a pair of stimuli are individually compared with reference values stored in memory to determine which of a given pair of values is the larger or which is the smaller (i.e., closest to the reference value). Reference values are context-dependent and are thought to represent the extreme values encountered in a given context. In our study, the extensive training on the values 1–9 may have prompted monkeys to use the value 1 as a reference point to determine which of two values was the smaller and 9 as a reference point to determine which of two values is the larger (33). Such an algorithm would produce the semantic congruity effect in response latency because the farther two values are from the reference point, the longer it takes to compare each of them to the reference point.

Regardless of which model best accounts for the semantic congruity effect, our demonstration of the effect in monkeys implicates the presence of a comparative process common to the comparison of Arabic numerals in humans and nonverbal numerical values in monkeys. Our data also indicate that the comparative process is not mediated by symbolic, linguistic, or otherwise uniquely human mechanisms. Instead, the semantic

[†]Examination of the choose larger function in Fig. 4 reveals that monkeys were faster at ordering large values, such as 8 vs. 9, compared with small values, such as 1 vs. 2. This pattern defies Weber's law and should be investigated further.

[‡]The 4-5 and 5-6 pairs were excluded because it was unclear whether they should be categorized as small or large.

congruity effect appears to be the signature of an evolutionarily primitive magnitude comparison mechanism.

We thank Herb Terrace for critical discussions about these data; and Kerry Jordan, Kerrie Lewis, Evan MacLean, Pierre Rojas, Jessica Ward, and all of the members of the E.M.B. laboratory for help in collecting data and discussing the results. This work was supported by a National Science Foundation Graduate Fellowship (to J.F.C.), National Institute of Child Health and Human Development Grant R01 HD49912 (to E.M.B.), a National Science Foundation Faculty Early Development Career award (to E.M.B.), and a Merck Scholars award (to E.M.B.).

1. Cantlon, J. F. & Brannon, E. M. (2006) *Psychol. Sci.*, in press.
2. Whalen, J., Gallistel, C. R. & Gelman, R. (1999) *Psychol. Sci.* **10,** 130–137.
3. Beran, M. (2001) *J. Comp. Psychol.* **115,** 181–191.
4. Brannon, E. M. & Terrace, H. S. (1998) *Science* **282,** 746–749.
5. Hauser, M. D., Tsao, F., Garcia, P. & Spelke, E. S. (2003) *Proc. Biol. Sci.* **270,** 1441–1446.
6. Nieder, A. & Miller, E. K. (2003) *Neuron* **34,** 149–157.
7. Gelman, R. & Gallistel, C. R. (2004) *Science* **306,** 441–443.
8. Barth, H., Kanwisher, N. & Spelke, E. (2003) *Cognition* **86,** 201–221.
9. Buckley, P. B. & Gillman, C. B. (1974) *J. Exp. Psychol.* **103,** 1131–1136.
10. Moyer, R. S. & Landauer, T. K. (1967) *Nature* **215,** 1519–1520.
11. Banks, W. P., White, H., Sturgill, W. & Mermelstein, R. (1983) *J. Exp. Psychol. Hum. Percept. Perform.* **9,** 580–582.
12. Shaki, S. & Algom, D. (2002). *Memory Cognit.* **30,** 3–17.
13. Holyoak, K. J. & Mah, W. A. (1982) *Cognit. Psychol.* **14,** 328–352.
14. Petrusic, W. M., Baranski, J. V. & Kennedy, R. (1998) *Memory Cognit.* **26,** 1041–1055.
15. Audley, R. J. & Wallis, C. P. (1964) *Br. J. Psychol.* **55,** 59–73.
16. Banks, W. P. & Flora, J. (1977) *J. Exp. Psychol. Hum. Percept. Perform.* **3,** 278–290.
17. Holyoak, K. J. & Walker, J. H. (1976). *J. Verbal Learn. Verbal Behav.* **15,** 287–299.
18. Moyer, R. S. & Bayer, R. H. (1976) *Cognit. Psychol.* **8,** 228–246.
19. Banks, W. P., Fujii, M. & Kayra-Stuart, F. (1976) *J. Exp. Psychol. Hum. Percept. Perform.* **2,** 435–447.
20. Holyoak, K. J. (1978) *Cognit. Psychol.* **10,** 203–243.
21. Cech, C. G. (1995) *J. Exp. Psychol. Learn. Mem. Cognit.* **21,** 1275–1288.
22. Smith, M. & Magee, L. (1980) *J. Exp. Psychol.* **109,** 373–392.
23. Sternberg, S. (1969) *Am. Sci.* **57,** 421–457.
24. Gallistel, C. R. & Gelman, R. (2000) *Trends Cognit. Sci.* **4,** 59–65.
25. Gordon, P. (2004) *Science* **306,** 496–499.
26. Pica, P., Lemer, C., Izard, V. & Dehaene, S. (2004) *Science* **306,** 499–503.
27. Choplin, J. M. & Hummel, J. E. (2002) *J. Exp. Psychol.* **131,** 270–286.
28. Cruse, D. (1976) *Lingua* **38,** 281–292.
29. Allan, K. (1986) *J. Semantics* **5,** 1–50.
30. Clark, H. H. (1969) *Psychol. Rev.* **76,** 387–404.
31. Gentner, D. & Goldin-Meadow, S., eds. (2003) *Language in Mind: Advances in the Study of Language and Thought* (MIT Press, Cambridge, MA).
32. Petrusic, W. M. (1992) *J. Exp. Psychol. Hum. Percept. Perform.* **18,** 962–986.
33. Brannon, E. M., Cantlon, J. F. & Terrace, H. S., *J. Exp. Psychol. Anim. Behav. Processes*, in press.