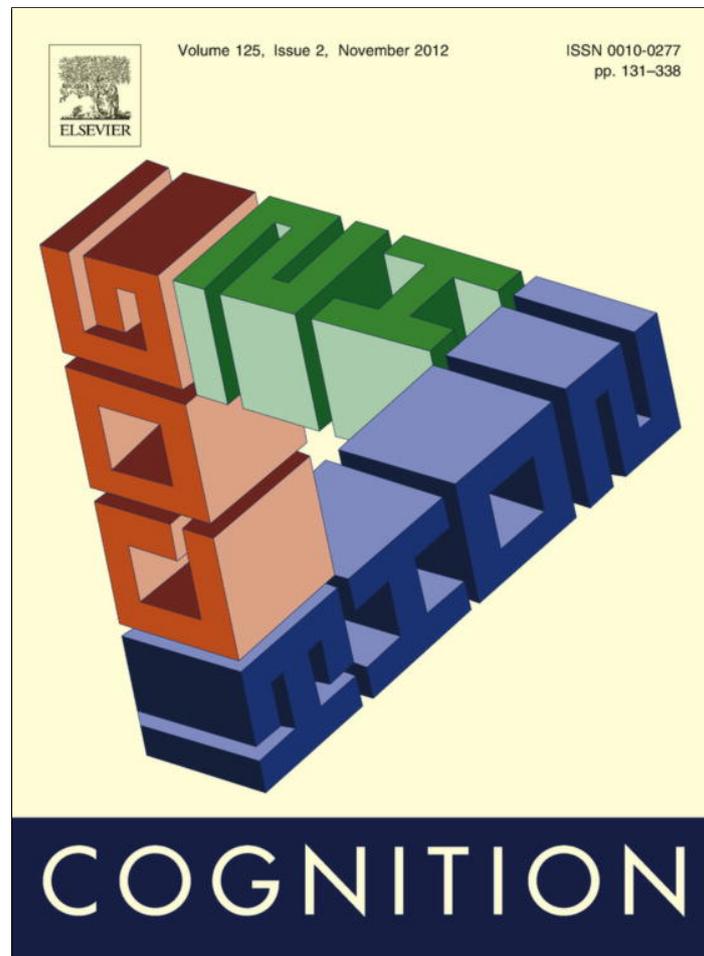


Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>

Contents lists available at [SciVerse ScienceDirect](#)

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Grey parrot number acquisition: The inference of cardinal value from ordinal position on the numeral list

Irene M. Pepperberg^{a,b,*}, Susan Carey^a^a Department of Psychology, Harvard University, United States^b Department of Psychology, Brandeis University, United States

ARTICLE INFO

Article history:

Received 4 September 2011

Revised 24 June 2012

Accepted 4 July 2012

Available online 9 August 2012

Keywords:

Animal cognition

Parrot cognition

Nonhuman numerical competence

Successor function

Parrot number acquisition

ABSTRACT

A Grey parrot (*Psittacus erithacus*) had previously been taught to use English count words (“one” through “six” [six]) to label sets of one to six individual items (Pepperberg, 1994). He had also been taught to use the same count words to label the Arabic numerals 1 through 6. Without training, he inferred the relationship between the Arabic numerals and the sets of objects (Pepperberg, 2006b). In the present study, he was then trained to label vocally the Arabic numerals 7 and 8 (“seven”, “eight”, respectively) and to order these Arabic numerals with respect to the numeral 6. He subsequently inferred the ordinality of 7 and 8 with respect to the smaller numerals and he inferred use of the appropriate label for the cardinal values of seven and eight items. These data suggest that he constructed the cardinal meanings of “seven” (“seven”) and “eight” from his knowledge of the cardinal meanings of one through six, together with the place of “seven” (“seven”) and “eight” in the ordered count list.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Nonhumans mostly draw on two distinct systems of mental symbols to represent number (see Carey (2009), Dehaene (2009), Gallistel and Gelman (2005) for reviews). Symbols in one system (the analog magnitude system) are mental magnitudes that are linear or logarithmic functions of the cardinal values of sets of items, and support computations of numerical equivalence and order, addition, subtraction, and ratios. But analog magnitude symbols only approximate the number of items in a set, are subject to Weber's law, and thus do not represent exact cardinality. A second system represents number only implicitly, with no mental symbols for cardinal values *per se*. Mental models in working memory are created for small sets of items,

with one symbol for each item; these models also support computations of numerical equivalence and order, and addition and subtraction, based on 1–1 correspondence. This parallel individuation system cannot capture, even implicitly, any number beyond working memory limits (~4).

To overcome limits of these evolved systems, humans invented representations using external symbols. Tally systems (notches on a stick or clay tablet, beads on a string) represent cardinal values of sets of individuals via 1–1 correspondence, using external symbols to transcend limits of working memory. The verbal numeral list, deployed according to Gallistel and Gelman's (1992) “counting principles” (CP), represents exact cardinal values for every numeral in the list. Counting principles state: numerals must be applied in order to items in a set to be enumerated, must be applied in 1–1 correspondence, and the last numeral in a count represents a set's cardinal value. No evidence exists for mental representations of exact cardinal values of sets >4 by adult humans in cultures

* Corresponding author at: Department of Psychology, Wm. James Hall, Harvard University, 33 Kirkland Street, Cambridge, MA 02138, United States. Tel.: +1 617 495 2852.

E-mail addresses: impepper@media.mit.edu, impepper@wjh.harvard.edu (I.M. Pepperberg).

lacking verbal numerals or a tally system (Frank, Everett, Fedorenko, & Gibson, 2008; Gordon, 2004; Pica & Lecompte, 2008), deaf adults lacking natural language input (Spaepen, Coppola, Spelke, Carey, & Goldin-Meadow, 2011), preverbal infants (e.g., Feigenson, Dehaene, & Spelke, 2004) or nonhumans (except, possibly, those taught explicit symbolic representations, see below).

CP acquisition is not easy (Carey, 2009; Fuson, 1988). For several years children assign cardinal meanings only to a subset of their known count list (Le Corre, Van de Walle, Brannon, & Carey, 2006; Wynn, 1990, 1992). Two-year-old English learners produce the count list, but assign exact cardinal meaning only to “one”; other numerals mean “some” or “plural”. Some 9 months later, they learn “two”; other numerals are “more than two”. A few months later they master “three”, then “four”. Only then do children *induce* the CP: that each successive numeral in the count list is exactly +1 than its predecessor. This induction separates them from subset-knowers (Sarnecka & Carey, 2008); they can now encode cardinal value expressed by any numeral in their count list (see Carey, 2009; Hurford, 1987; Klahr & Wallace, 1973 for general characterizations of the Quinian bootstrapping that underlies CP acquisition, as well as many other episodes of conceptual change [Carey, 2009]).

These facts raise two interrelated questions about nonhumans' number representations: (1) Are they *restricted* to analog magnitude representations and/or parallel individuation? (2) Is bootstrapping, dependent upon logical operations carried out on external symbols and the capacity to create mappings between distinct representational systems, uniquely human? If so, a new hypothesis would be introduced to (partially) account for humans' unique capacity for cultural knowledge construction. We explore these two questions by investigating whether a nonhuman can infer the cardinal value represented by a novel numerical symbol from its place in an ordered numeral list.

The bootstrapping process children use to perform this feat has several prerequisites. First, children must represent order in the numeral list. Second, they must represent cardinal values of the first numerals (“one, two, three, four”). Third, they must understand and represent that cardinal values of adjacent numeral pairs—“one–two”, “two–three”, “three–four”—each differ by exactly 1. Finally, they must induce that this pattern extends indefinitely—that cardinal values of *any* two adjacent numerals differ by 1. Nonhumans have demonstrated many, but not yet all, of these prerequisites.

First, nonhumans represent ordinal relations among arbitrary stimuli. Macaques can learn to order 3–7 random items (e.g., cup, tiger, chair, etc.; Chen, Swartz, & Terrace, 1997; Swartz, Chen, & Terrace, 2000; Terrace, Son, & Brannon, 2003). Chimpanzees (Matsuzawa, 2009), rhesus, and capuchins (e.g., Beran et al., 2008; Harris, Beran, & Washburn, 2007) acquire ordinal relations among Arabic digits “0” through “9” in the absence of knowledge of these symbols' cardinal values.

Second, animals can use the analog magnitude system, constrained by Weber's Law, to support numerical computations of more/less for sets of individuals (Dehaene, 2009).

For example, macaques (Brannon & Terrace, 1998, 2000) and pigeons (Scarf, Hayne, & Colombo, 2011) learned the rule “touch in order of increasing numerosity” for sets of diverse items in order “set size 1, set size 2, set size 3, set size 4”, generalized to new sets of 1–4 items, then to untrained sets of 5–9.

Third, nonhumans have mapped numerals to approximate quantities. Monkeys (e.g., Beran et al., 2008) learned to relate digits to corresponding sets of candies, but also possibly to hedonic value or reward probability. Later, taught two ordered lists—numerals or dot arrays rewarded by corresponding sets of candies, or one candy for choice of the larger array (Harris, Gullledge, Beran, & Washburn, 2010)—monkeys, without training, integrated two ordered lists trained separately in alternation, choosing the larger of two stimuli, one from each list (e.g., set of 5 dots, the digit “3”; see D'Amato and Colombo (1988) and Terrace et al. (2003) for similar data).

Finally, some nonhumans unequivocally map numerals to cardinal values of sets. Most notable are two apes, Matsuzawa's Ai and Boysen's Sheba, and our subject, a Grey parrot, Alex. The apes' training took years, proceeding piecemeal (Biro & Matsuzawa, 2001; Boysen, 1993; Boysen & Berntson, 1989; Boysen, Berntson, Shreyer, & Quigley, 1993; Matsuzawa, 1985; Matsuzawa, Itakura, & Tomonaga, 1991; Murofushi, 1997; Tomonaga & Matsuzawa, 2000). Thus, Ai first learned to touch the digit “1” when shown sets of one item versus two, then “2” was added to her symbolic repertoire, up to “9”. Ordinality had to be trained separately, as was equating an item set to a given digit. Sheba was taught similarly, to “8”. Children do not learn in this piecemeal fashion: Even “subset-knowers” order numerals of known cardinal value and facilely label sets with verbal numerals and construct sets to cardinal values specified by known numerals (Le Corre, Van de Walle, Brannon, & Carey, 2006). Eventually, apes mastered these abilities.

What representations supported these abilities? Human adults resolve analog magnitude representations in ratios of 7:8 or 8:9 (see Barth, Kanwisher, & Spelke, 2003; Halberda & Feigenson, 2008); possibly Ai, Sheba, and Alex mapped numerals to analog magnitudes. However, if enumerating sets under time pressure, Ai's reactions reflected a pattern of subitizing for sets of 1 to 3, then a time increase with additional items from 4 to 9, suggesting another mechanism. Sheba tapped items one at a time before touching the symbol for cardinality of that set. Such data are consistent with the suggestion that some representations drawing on 1–1 correspondence or counting underlie at least some of Ai's and Sheba's performance with large sets. But neither ape engaged in the bootstrapping process that underlies children's mastery of the counting principles. Neither showed savings in learning as successive numerals “5”, “6”, “7”, etc. were added to her repertoire.

The parrot, Alex, also took several years to learn to quantify sets of 1–6 items with vocal English labels (Pepperberg, 1987, 1994). But unlike children and apes, he did not acquire number labels in order, first learning “three” and “four”, then “five” and “two”, lastly “six”

(pronounced “sih”¹) and “one”. For each vocal label, he had to learn to manipulate his vocal tract in specific ways (i.e., simultaneously control muscles of his tongue, beak, larynx, glottis, syrinx, trachea; head position, possibly choana and esophagus; Patterson & Pepperberg, 1994, 1998), often a slow, difficult process (Pepperberg, 1999). A novel label with familiar phonemes could be acquired rapidly (even overnight; Pepperberg, 1983), but no initial number label fully fit that criterion. Thus, Alex *might* have quickly comprehended number labels learned later in training, his slow acquisition reflecting difficulty in vocal production. The present study separates these components of learning new verbal numerals.

Without further training, Alex’s number label *comprehension* matched his trained production, as in children (Le Corre et al., 2006; Wynn, 1990, 1992). Tests required identifying specific numerical subsets (e.g., “What color five?” for a 5 red cork subset intermixed with 4 green and 6 blue corks; Pepperberg & Gordon, 2005). Like Ai, he needed <1 s to quantify sets of 1, 2, 3, but much longer for 5 or 6, though >80% accuracy could be achieved given enough time (Pepperberg, 2006a). He also learned to use the same vocal labels for Arabic digits “1”–“6” *without any connections* to the relevant item set; then, *without training*, inferred the ordinality of his labels, tested by his stating the color of the larger *or* smaller Arabic digit in a paired set or “none” if they represented the same quantity (Pepperberg, 2006b). Thus, Alex provided evidence of an integrated system of number representations by verbal and Arabic numerals, representing cardinal values and their ordinal relations.

Alex, like Ai and Sheba, never induced the cardinal value of a given numeral from its place in the count list, that is, never engaged in the bootstrapping children use. He, like the apes, never learned the ordered list prior to mapping numerals to cardinal values. His novel number labels were never explicitly placed into a numeral list, though taught as numerals whose cardinal values differed by just 1 (“three” vs. “four”), or in one-more-than/one-less-than pairings with respect to known labels (e.g., “five” vs. “four”, “two” vs. “three” for the second set of numerals trained).

However, that no nonhuman has yet induced how the numeral list works does not prove they cannot. Here we tested the hypothesis that, were Alex taught labels for novel Arabic digits (the words “seven”, “eight”), and their ordinal position in the list (“six” is a smaller number than “seven” which is smaller than “eight”), he would infer the cardinal values of “7/seven” and “8/eight”. Training the two labels simultaneously avoided the possibility he would associate a novel label with any set of a previously unnamed cardinality. Training him to say “seven” (actually “sih–none”, as /v/ was difficult) and “eight” as labels for Arabic digits allowed us to separate the time needed to learn to produce the words from the process of learning what they meant. Finally, we assessed the crucial question: Did Alex infer cardinal meanings of “7/seven” and “8/

eight” from their positions in the ordered count list? Success by a parrot, whose last common ancestor with humans was 280 million years ago, would suggest that many of the logical and symbolic capacities that underlie the construction of an explicit symbolic representation of integers are not unique to humans, and perhaps are wide-spread in the animal world.

2. General method

2.1. Subject and housing

Alex, a male Grey parrot (*Psittacus erithacus*), 28 years old, had been the subject of cognitive and communicative studies for 27 years. For living conditions outside of testing and training, see Pepperberg and Wilkes (2004). Food and water were available at his vocal request. This study used several of his documented abilities: to learn new vocal labels, to label seven colors vocally (Pepperberg, 1999), to vocally respond to auditory queries on relative size/absence of size difference (trained initially for objects; Pepperberg & Brezinsky, 1991), to vocally label cardinal values of 1–6 item sets (Pepperberg, 1987, 1994), to use these vocal labels for Arabic digits “1”–“6” and to report ordinal relations between pairs of Arabic digits “1”–“6” (Pepperberg, 2006b).

2.2. Apparatus and materials

Arabic digits “7”, “8” were of various colors, wood or plastic, of the same sizes as “1”–“6” used previously (~2–3 cm; Pepperberg, 2006b). The felt-covered tray on which we put stimuli for all trials had been used in earlier studies: label comprehension (Pepperberg, 1990a, 1990b, 1992), object permanence (Pepperberg & Kozak, 1986), other number tasks (Pepperberg, 1994, 2006a, 2006b; Pepperberg & Gordon, 2005). It was concurrently used to train spatial concepts, phoneme recognition (Pepperberg, 2007), and test optical illusions (Pepperberg, Vicinay, & Cavanagh, 2008); it did not cue numerical tasks.

2.3. General training

Training was primarily via the Model/Rival (M/R) procedure (Pepperberg, 1981, 1990a). Briefly, it involves three-way *social* interactions among two humans and a bird. Typically, a bird sees two humans discussing one or more items in which it has already shown interest: A trainer presents, and queries another human about, the item(s) (e.g., “What’s here?”, “What color?”) and gives praise and the item(s) so as to reinforce the referential relationship between item(s) and label. The human queried may intentionally make errors (like the bird’s), and is punished by scolding and temporary removal of the item(s). Thus the second human is a model for a bird’s responses, its rival for a trainer’s attention, and illustrates effects of an error: S/he tries again or talks more clearly after an incorrect or garbled response, demonstrating corrective feedback, as well as the reason for learning specific sounds of the label. A bird is included in interactions and rewarded for succes-

¹ All number labels were standard English except for “six,” which he pronounced “sih,” the final /s/ being difficult to produce (Patterson & Pepperberg, 1998)

sive approximations to a correct response; training is thereby adjusted to its level. Roles of human trainer and model are also reversed to emphasize that one is not always the questioner and the other the respondent.

This system was used to train Alex to label Arabic “7”, “8” with vocal numerals. As he used “sih” for six, and first used “sih–one”, then “sih–none” for “seven” (/v/ being difficult), he was now taught “sih–sih” for six, so we did not cue “seven” (“sih–none”) by waiting for a second syllable. “Eight” was not distinguishable from human speech. “Sih–one/sih–none” and “eight” were trained with Arabic digits exclusively; “sih–sih” was (re-)trained for both 6-item sets and digits as Alex had previously inferred the connection between a 6-item set and the symbol “6” (Pepperberg, 2006b). Training queries for “6” were “What number?” and for “7” and “8” primarily “What’s here?”, although rarely “What number?” was also used for these digits, providing some evidence that “7” and “8” were also numerals.

Next, via M/R training, Alex learned that, for Arabic numerals, “6” < “7”, “7” < “8”, “6” < “8”, “8” > “7”, “8” > “6”, “7” > “6”. We put two digits of different colors on the tray, modeling replies to “What color number (is) bigger/smaller?”, to teach ordinal relations among “6” (“sih–sih”), “7” (“sih–none”), and “8” (“eight”), and provide unequivocal evidence for the numeral status of “7” and “8”.

2.4. Testing

Alex received several tests. First, a pre-training exam explored how Alex would label sets of 7 and 8 items prior to training, particularly in relation to the meaning he had assigned to “sih”/“six” in his production/comprehension vocabulary. If “sih/six” meant approximately six, or “ $n > 5$ ”, he might be expected to apply it to sets of 7 and 8, as they have cardinal values for which he had not been taught a label. Second, after each round of training, he received Training Efficacy Exams: Exam 1 to see if he could label the digits “6”, “7”, “8” as “sih–sih”, “sih–none”, “eight”; Exam 2 to see if he had learned ordinal relations among “6”, “7”, “8”. A List Integration Exam then tested if he had integrated ordinal positions of “7” and “8” with numerals “1”–“5”, the latter never being involved in any current training. Finally, crucial Cardinal Induction Tests (production, comprehension) examined if, without any training, he produced “sih–sih”, “sih–none” and “eight” as labels for 6-, 7- and 8-item sets, respectively, and if he comprehended “six”, “seven”, and “eight” as labels for sets of these cardinalities.

2.4.1. Criteria prior to testing

The pre-training exam was, by definition, exploratory with respect to existent knowledge.

For tests on Arabic numeral labeling (Training Efficacy Exam 1) to begin, Alex’s “sih–sih”, “sih–none”, and “eight” had to be recognized in blind trials with >90% inter-observer agreement, a criterion based solely on speech clarity: Thus, he could misidentify a digit, but if both presenter

and a listener, blind to the digit Alex was labeling, agreed on what they heard, the trial counted toward the 90% criterion. Exam 1 then required 90% labeling accuracy before training of ordinal relations among “6”, “7”, and “8” began. Had Alex failed, he would have received more training and retesting.

Although Alex used “sih–sih” for “6” and “sih–none” for “7”, humans used standard “six” and “seven” in querying him, and he replied appropriately. Thus, despite any possible production issues, Alex knew “sih–sih” and “sih–none” corresponded, respectively, to “six” and “seven”.

Formal testing on ordinality of Arabic numerals “6, 7, 8” (Training Efficacy Exam 2) began after Alex was correct >70% for ordinal training trials for two consecutive sessions. If ordinal test accuracy on the trained set (“6”, “7”, “8”) was $\geq 80\%$, formal testing (List Integration Exam) could begin on the *untrained* ordinality of Arabic numerals “1–5” with “7, 8”.

Cardinal Induction Tests, which examined Alex’s inference of what labels should apply to 7- and 8-item sets, could begin only after he had demonstrated, via Training Efficacy Exams 1, 2 and the List Integration Exam, that he knew the verbal labels for the Arabic numerals and order of “7/sih–none” and “8/eight” with respect to the other verbal and Arabic numerals in his repertoire.

2.4.2. General test procedures

Testing occurred 1–4 times/wk, excluding breaks (holidays, intersession, other examiner absences, periods of noncompliance), from December, 2005 to April, 2007. Training Efficacy Test 1 occurred December, 2005 to February, 2006; Test 2 occurred March, 2006; the List Integration Test occurred April, 2006; and Cardinal Induction Testing (production, comprehension) occurred May to June, 2006, with four extra test trials on production in April, 2007. Exam queries were presented intermittently either during free periods (when Alex was not being trained on any topic) or during sessions on unrelated topics (e.g., using him to train another bird on colors; during tests on optical illusions) until all exam queries were presented. Testing never occurred under time constraints.

Noncompliance meant just that—sessions clearly unlike the norm, in which Alex neither maintained his gaze on the tray nor made eye contact with the tester, but instead endlessly preened, requested to be returned to his cage or requested treats he then discarded, or uttered strings of irrelevant labels (e.g., colors not on the tray and thus not possible response choices). As whenever he acted thusly (Pepperberg, 1992; Pepperberg & Gordon, 2005; Pepperberg & Lynn, 2000), we ended the session. He received a brief time-out; trials were duly noted and repeated in later sessions. Such behavior never occurred in Training Efficacy Exam 1 nor Cardinal Induction Tests, but for 1/15 trials in Training Efficacy Exam 2 and 1/26 trials in the List Integration Test. Counting those two trials as errors would not affect the overall pattern of data presented below.

Test protocols, including procedural details, precautions against inadvertent and expectation cuing, are in Pepperberg (1981, 1990a, 1994); summaries are below. If he said

the correct label in a test trial, he received praise and the item to which the query referred, or could request an alternative reward. No further repetition of the trial occurred (i.e., there was only a single, “first trial” response). After an incorrect or indistinct reply, the examiner removed the tray of objects, turned his/her head, and emphatically said “No!” and repeated the query (Pepperberg, 1981, 1987). The procedure penalized both a “win-stay” strategy and noncompliance;² presentation continued until a correct response was made or four attempts occurred; errors were recorded. However, Alex rarely erred (we report first trial responses) and, most importantly, errors almost always involved perceptual issues (e.g., color vision) or repetitions of parts of our queries, and only by querying him again could we sort out such issues from those involved in knowledge of the task being tested. See “Scoring” section below. He was not corrected for numerically wrong responses by being tested until he gave the right answer, as only once did we repeat a trial on which he had made a numerical error; thus these repeated trials during testing did not serve as training trials.

As in all Alex's studies, protocol differed from those used with other animals in two respects. (1) Tasks capitalized on Alex's use of the vocal mode. (2) Each trial was presented intermittently during training and testing of other unrelated topics. Alex's responses thus had to be chosen from his entire repertoire (>100 vocalizations, including labels for objects, locations, materials, relations, colors, shapes, foods, etc.) and from among numerous topics concerning various exemplars and queries in each session; that is, Alex had to attend to the specific type of query posed. This design increased task complexity and prevented several forms of cuing (see below; Pepperberg, 1999; Premack, 1976).

2.4.3. Controls against expectation and experimenter-induced cuing

Tests included specific precautions to avoid cuing. One control was a design such that each test session was presented intermittently during free periods or work on unrelated topics. A tester posing series of similar queries may come to expect a particular answer and unconsciously accept an indistinct (and by our criteria incorrect) response of, for example, “gree” (a mix of “green”/“three”) for “green”. As noted above, Alex's responses had to be chosen from his entire repertoire and from among numerous possible topics during each session; each session had only one number array. Second, a human other than the one testing digits or sets on the tray (one of five possible people), who did not know what was shown, confirmed the answer; his/her interpretation of Alex's response was thus not likely influenced by expectation of a certain number label. Only after confirmation was Alex rewarded (Pepperberg, 1981). Third, this evaluator was not likely influenced by hearing the type of query posed: In a prior study, transcriptions of contextless tapes of Alex's replies in a session

² With respect to noncompliance: In earlier studies, Alex, caring little for any reward, realized he could quickly get through a test no matter what he said. After learning his trainers would persevere, he realized he had to respond correctly for a test to end.

agreed 98.2% with original evaluations (Pepperberg, 1992).³ Fourth, Alex had not been trained to quantify 7- or 8-item sets, thus no overlap occurred between training and final critical tests. For tests on number labels and ordinality, persons acting as trainers never performed tests; thus Alex could not have picked up on trainer-induced cues specific to a given label (Pepperberg, 1981). Moreover, six different humans (total) were involved in testing, thus presence of a particular person could not cue a number session. Fifth, several different common items were used to test set cardinality, including those for training the other birds (e.g., variously sized jelly beans for training on colors); thus particular items would not cue that a number test was in progress. Controls for mass/ contour/size were in place (Section 2.4.4). Two Arabic numerals might cue an ordinality test, but the data obtained would confirm whether the correct ordinal relations had been acquired.

Roughly 10% of trials on identification of collections were videotaped to check further for cuing and interobserver reliability. Video was used to check on Alex's attention and motivation. Observers for reliability had never heard Alex speak and did not watch the video but listened as if scoring in the laboratory, stating what they heard him say. If interobserver reliability with blind coders was high, we could be assured of the validity of other trials (results in Section 4.5). Given that Alex was far more interested in the human who was taping, the camera and mike, than in the numerals and sets to be enumerated, such trials were difficult to execute and kept to a minimum.

2.4.4. Test materials

All items varied in size/contour/mass *across* and, except for blocks (1" cubes), *within* trials (i.e., different items with varying size/contour/mass/density were used across trials). “Sticks” could be 1/4" × 2", 3/8" × 4.5" and 3/4" × 6"; jelly beans varied among “Jelly Belly” and standard sizes, candy hearts were at least two sizes, woolen pompons could be 1/2", 3/4" and 1" in diameter, corks and paper were re-used and chewed by Alex to various sizes and shapes, “treat” was a somewhat squarish biscuit, whose edges could vary from 1.7 cm to 2.5 cm. Thus Alex could not be responding based on variables other than number. For instance, seven jelly beans would have different total volume and contour on different trials, be displayed in different configurations on different trials, and would have vastly different total volume, contour, and density than seven blocks or pompons.

2.5. Scoring

We performed statistical analyses on the percentage of first trial numerically appropriate responses. Three response types that did not literally answer a query were

³ This percentage represented 106 matches of 108 vocalizations. As an additional control, the principal investigator made two transcriptions of a student, new to the lab, as he responded to the same questions as Alex. The first transcription was live; the second, made several days later, was of a tape from which all questions had been edited. The two transcriptions of the student's vocalizations matched to within 95.8% (68 of 71 vocalizations).

not counted as errors: (1) Alex's reply to, for example, "What color number (is) bigger?", with the appropriate Arabic numeral label rather than its color was accepted, as it was numerically correct. We did repeat the query to force him to attend to the task, but such repetition clearly did not teach him the numerically correct response. (2) A perceptual color confusion was not considered an error—such confusion stems from parrots' splitting color categories, such as orange, differently than do humans (Bowmaker, Heath, Das, & Hunt, 1994, 1996; Pepperberg, 1994, 2006a, 2006b).⁴ If he said "yellow" to, for example, a choice between orange and green, we provided the opportunity to self-correct, which he always did; had he switched to the incorrect alternative, an error would have been recorded. (3) If he first repeated part of a query before responding appropriately, such behavior was not discouraged to maintain vocal interaction, but the query was repeated (up to four times) until he gave an actual answer. Thus, the first numerically relevant response, correct or not, provides the data analyzed below. However, tables provide a complete record of all Alex's responses on Pretest, Training Efficacy, List Integration, and Cardinal Induction Exams—ALL data collected probing comprehension/production of "sih-sih/six" "sih-none/seven" and "eight".

Statistical analyses primarily were binomial tests, with various chance levels. We could have used the probability of Alex producing any of his ~100 utterances (all labels were always available from which to choose). Instead, we assumed he would always attend and respond numerically relevantly to queries of "How many X?", "What number?", "What color number bigger/smaller?", and "What color 'N'?" (i.e., not respond randomly to the task). He gave <1% off-topic single labels responses in testing (1/103 trials). Chance could have been considered 1/8, as if he randomly guessed among all number labels he could produce after hearing "How many X?" We more conservatively used chance of 1/3, as if he were choosing to respond with either of his two new labels "sih-none" and "eight", or "sih-sih", all labels focused on during training. For ordinality tests, chance was set at 1/2.

3. Procedures for tests

3.1. Pretests: Application of "sih/six" to 6-, 7-, and 8-item sets; production and comprehension

Before any training to label Arabic digits "7", "8", Alex had pretests (early October, 2004) to examine if he might apply "sih" to any set >5 or approximately 6. He saw, with no time constraints, in individual trials, 7, 8, or 9 items, two trials/set size (random order, interspersed with trials on

smaller sets and non-number tasks). We asked "How many X?" He was neither rewarded nor scolded whatever his reply, simply told "OK"; we then went to the next query. In a comprehension pretest (6 trials total), he saw four trays with sets of various numbers of items, including 7 or 8 but omitting 6 (e.g., 3 yellow wool, 4 blue wool, 7 green wool), and was asked "What color six?" He was also asked about two sets for which he knew numerals (3-, 5-item sets) to ensure he did not acquire a bias to respond "none".

3.2. Training Efficacy Exam 1: Labeling of Arabic numerals

Various Arabic numerals were placed on the tray; Alex was asked "What number?" He had 10 trials for each newly trained digit ("7, 8") and for "6", because he had learned a new two-syllable label for "6". Number and non-number queries (e.g., "What color key?", "What toy?") were mixed roughly equally to keep him from focusing on number labels exclusively. Likewise, one trial for each prior learned numeral ("1"–"5"; 8% of total trials) was inserted so he could not focus on the newer labels. Data were not collected on these five trials (too few for statistics) and he had already extensively verified his knowledge of these digits (e.g., Pepperberg & Gordon, 2005); he made few if any errors on these numerals and was 100% on objects/colors/shapes/materials. Using the same tray, we tested optical illusions simultaneously, for which color labels were correct answers (Pepperberg et al., 2008).

3.3. Training Efficacy Exam 2: Ordinal relations among numerals

After Alex showed he could label Arabic digits "7", "8", and been trained on "6" < "7", "7" < "8", "6" < "8", he was formally tested on these as well as on identical pairings. For each pairing (e.g., "6", "8"), two trials asked the color of the bigger numeral and two the color of the smaller, with right-left placement of the larger or smaller numeral counterbalanced across trials, for a total of 12 trials. One trial probed which numeral was bigger or smaller for each of "6" vs. "6", "7" vs. "7" and "8" vs. "8". Tests were performed as in Pepperberg (2006b), simultaneously with optical illusions tests.

3.4. List Integration Exam: Novel ordinality pairings

We tested if Alex could, without training, integrate the order of "7", "8" with his previously inferred ordinal ranking of "1"–"6" (Pepperberg, 2006b). Tests thus paired "7", "8" each with "1"–"6" and with each other. Again, for each comparison, one trial asked the color of the bigger numeral; the other, the color of the smaller, for a total of 26 trials. As right-left placement was irrelevant in Training Efficacy Exam 2, placement was now randomized. Tests were performed as in Pepperberg (2006b) and, again, with tests on optical illusions.

⁴ Wooden numerals—the initial stimuli—were always standardized with non-toxic paints for which Alex learned to use a specific color label; however, he chewed these numerals as his reward and they could not be replaced. We thus had to use magnetized plastic "refrigerator letters" for this study, and color errors occurred with these numerals. Paint that sticks to plastic is toxic, and because Alex is given the numeral to chew as his initial reward, could not be used. Note he did not always err on such items, and would answer correctly when asked a second time, suggesting that the color clearly was on a boundary (e.g., if not yellow or red, it had to be orange).

3.5. Induction of the cardinal meaning of “7/sih–none” and “8/eight”

3.5.1. Numeral production

We next tested if Alex had inferred the connection between the Arabic numerals and sets of 7 and 8 items. Sets ($n = 30$) of 7 or 8 various items were placed in random arrays on the tray; he was asked “How many X?” To avoid focus on the new labels, he was given a few (2–3) trials with smaller numbers of items and three trials with six items. Data on these two/three trials of <6 items went untallied, as no statistics could be performed, but he was generally correct.

For 13 of 33 total trials on 6, 7, 8, we added or removed 1 or 2 items to or from the previous set as he watched, then queried him again. These probes served two purposes. (1) Had Alex made the inference being tested, adding/subtracting 1 or 2 might focus him on number. (2) Had he inferred the relation between ordinal position and cardinality, he could work out the cardinal value of a set so transformed, had he been correct originally. To ensure adding/subtracting did not signal an error, six adding/subtracting trials adjusted a set to which he had assigned the correct cardinality; the other seven adjusted a set on which he erred. All addition/subtraction trials are noted in Table 5 (Section 4).

3.5.2. Numerical comprehension

As in pretests, three sets of varying cardinalities were intermingled on the tray, only now up to 10 items could be in a set. He was asked the color of a targeted subset (e.g., “What color seven?”) for 12 trials. On one trial, he was asked about a set of four items to avoid focus on new numbers. Comprehension trials, begun after production trials ended, were administered mixed with with optical illusion tests (Pepperberg et al., 2008). Unless noted, item sizes in tests varied within/across trials.

4. Results

We report Alex’s data in five parts, corresponding to the five tests described above.

Table 1

Baseline production data for collections of items (blocks, wool, etc.) prior to any training^a.

Number Presented	Responses
7 Treats	No initial response, question repeated 3×; eventually said “sih”
9 Beans	No initial response, question repeated 3×; eventually said “sih”
8 Wool	“sih” after much hesitation
9 Corks	“sih” after much hesitation
8 Blocks	No initial response, question repeated 2×; eventually said “sih”
7 Corks	No initial response, question repeated 3×; eventually said “sih”

^a Items differed in size across trials; except for blocks, items were differently-sized w/in trials.

Table 2

Baseline comprehension data for collections of objects prior to any training; absent six^a.

Trial type	Question	Response
4 Blue sticks, 5 orange sticks, 9 green sticks	What color 6?	None (first said “sih”)
3 Yellow blocks, 4 blue blocks, 7 green blocks	What color 6? What color 3?	None Yellow
5 Grey beans, 7 blue beans, 8 yellow beans	What color 6?	None (first said “sih”)
5 Green wool, 8 purple wool, 9 rose wool	What color 6? What color 5?	None Green

^a Items differed in size across trials; except for blocks, items were differently-sized w/in trials.

4.1. Pretests: The application of “sih/six” to sets of 7, 8 and 9 prior to training

Presented with 7, 8, or 9 items and asked (e.g., “How many cork?”) to label cardinality of a set for which he as yet had no verbal numeral, he initially refused to reply on 4/6 trials (first two, last two); he sat quietly after seeing the set or asked to return to his cage. As expected, he had no knowledge of labels for 7 or 8, lacking training. Only when badgered and forced to respond did he use the label representing his largest currently trained quantity (“sih”; Table 1). The data could imply “sih” meant “about 6” or “set of cardinality >5”, but his actions suggest he knew his standard number labels were incorrect; he did not, as when “bored” (Pepperberg, 1992; Pepperberg & Gordon, 2005), give strings of wrong answers, requests for treats, or turn his back and preen. His hesitancy to respond resembled that in his earlier addition study, when asked to label a total absence of items (Pepperberg, 2006a).

Comprehension pretest data suggest that “six” (a) did not mean “cardinality >5” to Alex, and (b) that his representation of six was precise enough to distinguish sets of 6 from sets of cardinalities of 4, 5, 7, 8 and 9. Asked “What color six?” for the (absent) 6-item set on trays with various numbers of differently colored items, including sets of 7 and 8, Alex replied “none” on all four trials, though twice he first repeated the label in our query—“sih” (Table 2). Asked on two trials for colors of sets that were present, to ensure he was attending to the stimuli, he replied correctly (Table 2).

4.2. Training Efficacy Exam 1: Labeling of newly-trained Arabic numerals

Alex was tested on referential labeling of digits 6, 7, 8 only after blind observers agreed what vocal label he was producing (whether or not it was correctly used). He needed over 13 months of training to consistently produce “sih–none” (seven) and replace “sih” with “sih–sih”; “eight”, part of his existent label “grate”, appeared in his first training session on that numeral. He had two to four number sessions/wk from October 18, 2004 through

November, 2005 with breaks for scheduled term vacations. Each session averaged two trials for each numeral (min. one, max. four), and each trial could have one to five M/R demonstrations (more modeling early in training, less at the end).

We did not transcribe all training trials; a sample subset suggests Alex acquired each new label differently. For 6 he not surprisingly initially said “sih”, which had been accepted for >15 years. If 7 preceded 8 in a session, he initially often said its color or, after receiving modeling, “s...n”, later shifting to “s...one” or “s...none”. He said “eight” to 8 over half the time, sometimes saying “T” or “ay”; he might confuse it with “three” or state its color. If 8 preceded 7, and he correctly labeled 8 and was rewarded, he persistently used “eight” for 7 in that session early in training. He was disinterested in some sessions, giving strings of wrong answers, carefully avoiding correct replies (Pepperberg & Gordon, 2005), which might explain his slowness in learning.⁵

In December, 2005, after clarity criteria were met, formal tests began on labeling Arabic digits. Alex's score was 10/10 for “6” and “7”, and 8/10 for “8” (calling it “sih–none” once and “three” once; Table 3). He never replied with off-topic labels or by repetitively listing incorrect responses.

In sum, during the year Alex took to learn to say “sih–none” and “eight” he also learned which Arabic digit—“7” or “8”—each labeled. He also successfully learned to label “6” as “sih–sih”.

4.3. Training Efficacy Exam 2: Ordinal relations between “7” and “8”

Alex learned ordinality for “6”, “7”, “8” faster than he learned their labels. Training began mid-December, 2005, after he scored >90% on the first half of testing on Arabic numeral labels. Ordinality sessions occurred about once/wk until the end of February, 2006, when he began replying correctly on a majority of trials. Sessions averaged one or two trials/pairing, one to five M/R demonstrations/trial (more early in training, fewer or none later). On Training Efficacy Exam 2 (March, 2006), which tested ordinality for “6”, “7”, “8”, including identical pairings (e.g., “7” vs. “7”, for which the reply was “none”), Alex scored 13/15 (86.7%, $p < 0.01$, binomial test, chance of $\frac{1}{2}$). He erred on one “6” vs. “7” and one “6” vs. “8” trial.⁶ He had learned ordinal relations among Arabic digits “6”, “7”, “8”.

⁵ Previously, if he uttered a novel label during sound play and received an interesting item in response (e.g., desirable food for “grain”, a paper clip “chain”), he quickly connected label and item; if the item was not of interest (dried banana chips for “banacker” = “banana” + “cracker”), he dropped the label (Pepperberg, 1990b). Such disinterest (i.e., to some piece of plastic or wood differing only slightly from familiar 1s and 3s) might explain his initial lack of object-label correlation and early use of “eight” in response to 7s and 8s.

⁶ These data did not count one noncompliance trial, when he started to list every color not on the tray, as an error. If that trial is counted, his score was 12/15 (80%, $p = 0.02$). Alternatively, one could argue that chance should be $\frac{1}{3}$, for there were three responses relevant in this task (color of one Arabic numeral, color of the other, and “none”). With chance of $\frac{1}{3}$, the probability of 12/15 correct is < 0.001 .

4.4. List Integration Exam: Novel ordinality pairings

Alex now had to *infer* ordinality of “7”, “8” with respect to *all* numerals; that is, report “What color bigger” and “What color smaller?” for comparison of all numerals “1”–“6” with “7”, “8”. We also repeated tests of “7” vs. “8”. Alex scored 21/26 (Table 4; 81%; $p < 0.01$, binomial test, chance of $\frac{1}{2}$). For four of the five errors, he initially said “none” (for “7–2”, “6–7”, “3–8”, “6–8;” order is of numerals on the tray). Once (in 26 trials) he named the color of the larger (“7”) rather than the smaller (“4”) numeral when asked “what color smaller?” (larger was orange, he said “rose”).⁷

Alex scored 17/20 correct on comparisons of “7” and “8” with “1”–“5” ($p = 0.001$, chance of $\frac{1}{2}$); scores did not differ (Fisher's exact test, $p = 0.56$, two-tailed) from those among “6”, “7”, and “8” (4/6 correct), though the latter pairings had been trained. Alex's had thus inferred the exact order of the numerals from “1”–“8”. Mastery of this ordinal list was a prerequisite for our central question: Would he infer cardinal values of “sih–none” and “eight” from his knowledge of the cardinal values of “1/one” through “6/sih–sih” and knowledge of ordinal relations among his symbols for “1”–“8”?

4.5. Inferring the cardinal meanings of “seven (sih–none)” and “eight”

4.5.1. Production

Every production trial is listed in Table 5 in the order administered. Alex correctly labeled the very first 7-item presentation, then was correct on first trials for the next two presentations of 7 items; on the first 8-item presentation he said “sih–one”, then self-corrected (no human intervention) to say “eight”. He was then correct on first trials of the next five 8-item presentations. Other than trials 1 through 3 (sequentially administered during BBC filming)⁸ or sequential trials with added/subtracted items, presentations were not massed, but administered individually, usually 1 day apart. Alex's only responses were “sih”, “sih–none” and “eight”, so we used chance of $\frac{1}{3}$ (drawing from these three responses). On first trials for 7- and 8-item sets, he scored 22/30⁹ ($p < 0.001$). For sets of 7, he scored 11/

⁷ As described in Scoring, two types of responses that did not answer a question directly were not counted as errors. On two trials, he labeled the orange numeral (the numerically correct response) initially as either yellow or rose, and corrected this to “orange” when asked again. There were no yellow or rose objects on the tray; these responses reflect color confusions traceable to parrot color vision. On a “6” vs. “8” trial, he replied “sih–sih” to “What color smaller?”—correct numerically—but not a color, which he gave when asked again. One non-compliance trial was aborted and repeated, the response from this repeated trial contributing to the data. This aborted trial was not tabulated in the data. Even if it is included as an error, his score, 20/26, is above chance. And, if we take $\frac{1}{3}$ as chance, as his responses were either colors of the larger or smaller digits or “none,” the probability of 20/26 trials correct is < 0.001 .

⁸ BBC filming took an entire day, in which the first hour or two involved his habituation to the crew, cameras, etc. Thus, by the time we filmed these trials, several hours had elapsed.

⁹ We planned to present fewer trials for each number, but the tape containing our interobserver reliability trials de-gaussed, and we had to run two more trials.

Table 3
Arabic numeral identification.

Numeral	Response	%	Probability (chance $\frac{1}{3}$) ^a
6	“sih-sih” (10×, first trial)	100	<0.01
7	“sih-none” (10×, first trial)	100	<0.01
8	“eight” (8×, first trial; 2× second trial: errors of “sih-none”, “three”)	80	<0.01

^a Even if we use chance as $\frac{1}{2}$, as if he was choosing only between “sih-none” and “eight”, ignoring all other number labels, his overall score (28/30) is statistically significant ($p < 0.001$) and, individually, ≤ 0.05 .

Table 4
Summary of results and errors for all ordinality trials.

Description of trials	Score	%	Probability	Error description
Arabic numerals, every numeral–numeral trial (number irrelevant errors not included)	21/26	81	<0.01	Said “none” (4×); wrong choice (1×)
Arabic numerals, every numeral–numeral trial (number irrelevant errors included)	18/26	69	0.038	Said number, not color (1×); said “none” (4×); yellow–orange confound (2×); wrong choice (1×)

16 ($p < 0.01$). For sets of 8, he scored 11/14 ($p < 0.01$). These are conservative tests; he actually had 8 numerals to draw from in response to a “how many” query.¹⁰ For 6 items, he scored 3/3.

In sum, with no training on cardinal meanings of “sih-none” and “eight”, Alex labeled 7- or 8-item sets above chance levels. Further, scores likely underestimate his ability. On one trial, counted wrong (#18), he self-corrected after being told “no” but before he could be queried again. He was also correct on 11/13 trials where 1 or 2 items were added/removed from a set (not including the self-corrected response on #18), suggesting that adding/subtracting focused him on number and, perhaps, that he had an implicit representation of the successor function. For example, when he had correctly labeled a set as “eight” and 2 were removed, he said “sih-sih” with no hesitation.

4.5.2. Comprehension

Alex was simultaneously presented with three sets of items, each set differing in color and quantity (range, 2–10) and asked “What color N?” (N = “seven” or “eight”, plus one control query of “four”). Comprehension data show he mapped “seven” and “eight” to sets of 7 and 8, respectively (Table 6). Indeed, comprehension surpassed production, as was true for smaller quantities (Pepperberg, 1994; Pepperberg & Gordon, 2005). Overall, he scored 12/12 (chance of $\frac{1}{2}$, $p < 0.01$): 7/7 for “seven” ($p < 0.01$) and 5/5 for “eight” ($p < 0.01$). On one trial, he repeated the label “eight” four times before correctly providing the color of the 8-item set; we did not count such repetitions as errors (see above). On another trial, he said “sih-sih” and immediately gave the color of the 6-item set; asked again “What color eight?”, he immediately responded correctly. He also responded correctly when asked the color of a smaller quantity (control trial, “4”). Thus, lacking training to connect labels “sih-none” and “eight” with 7- or 8-item sets,

Alex inferred the connection. Note, he did not use “sih-none” or “eight” for the largest sets available (e.g., 9, 10) which would be expected had “sih-none” or “eight” meant anything simply >6 . These data show his inferred representations of cardinalities expressed by “sih-none” and “eight” were precise enough to distinguish 7 from 6, 8, 9 or 10, and 8 from 6, 7, 9, or 10.

4.6. Interobserver reliability—external evaluators unfamiliar with Alex and the experiment

Responses on all trials were coded on-line by an experimenter other than the one who asked the question, and this response identifier was blind to the array Alex was observing when queried. As a further test of Alex’s production, four naïve observers heard audio portions of tapes of 6 trials. Two agreed 100% with what the experimenter had judged Alex to say, two missed a color label (“What color eight?”, identifying Alex’s very soft “green” as “grey”; nothing grey was on the tray). No evaluator missed Alex’s “sih-none” or “eight”. Thus, agreement was 22/24 (91.7%).

5. General discussion

Humans in numerate cultures represent the cardinality of a set by the numeral whose ordinal position in a count list determines its cardinal value. Their children require 1½–2 years of experience with number before mastering this culturally constructed representational system. Here we asked if a nonhuman could acquire such a system; the answer appears to be yes.

Alex’s taking over a year to label Arabic digits 7 and 8 as “sih-none” and “eight”, and to assign a new label “sih-sih” to “6”/6 items, likely reflected difficulty in vocal production. Unlike children, who typically need few trials to acquire new lexical entries, Alex’s protracted process involved learning to configure his vocal tract to produce specific human phonemes (Patterson & Pepperberg, 1994, 1998). However, after mastering the new lexical forms, and mapping them to respective Arabic digits, he required only a few months, with roughly one

¹⁰ Even if we were to be extremely conservative and treat chance as $\frac{1}{2}$ (i.e., assuming responses only of “7” or “8,” even though some questions required a response of “6”), for first trials for sets of 7 and 8 items, $p < 0.01$; for judging sets of 7, $p = 0.07$, and for sets of eight items, $p = 0.02$.

Table 5

Production data for collections of objects; question are “How many X?”^a.

Trial	Quantity	Results
1	7 Mixed treats	Alex says “sih–none”
2	6 Mixed treats	(One taken away from set 1) Alex says “sih–sih”
3	8 Mixed treats	(Two are added to set 2) Alex says “sih–none”, self-corrects to “eight”
4	7 Mixed beans	Alex says “sih–none”
5	7 Small beans	Alex says “sih–none”
6	8 Mixed treats	Alex says “eight”
7	7 1” Blocks	Alex says “eight”
8	8 1” Blocks	(One added to set 7) Alex says “eight”
9	6 1” Blocks	(Two removed from set 8) Alex says “sih–sih”
10	7 Mixed keys	Alex says “sih–sih”
11	6 Mixed keys	(One is removed from set 10) Alex says “sih–sih”
12	7 Mixed keys	(One is added to set 11) Alex says “sih–none”
13	8 Mixed corks	Alex says “eight”
14	7 Mixed wools	Alex says “sih–none”
15	8 Mixed sticks	Alex says “eight”
16	7 Mixed sticks	(One item removed from set 15) Alex says “sih–none”
17	8 Mixed wheats	Alex says “eight”
18	7 Mixed wheats	(One item removed from set 17) Alex says “eight”
19	7 Mixed wheats	(One item added and immediately removed from set 19), Alex says “sih–none”
20	7 Mixed treats	Alex says “sih–sih”
21	7 Mixed treats	(One item removed and immediately added to set 20) Alex says “sih–none”
22	8 Mixed wheats	Alex says “sih–none”
23	8 Mixed wheats	(One item removed and immediately added to set 22) Alex says “eight”
24	8 Mixed beans	Alex says “sih–sih”
25	8 Mixed beans	(Two removed and immediately added to set 24) Alex says “eight”
26	7 Mixed spools	7 orange spool, one blue spool, Alex asked “How many orange spool?”; he says “sih–none”
27	8 Mixed hearts	Alex says “eight”
28	7 Mixed hearts	Alex says “eight”
29	7 Mixed hearts	(One item removed and immediately added to set 28) Alex says “sih–none”
30	8 Mixed beans	Alex says “eight”
31	8 1” Blocks	Alex says “eight”
32	8 Mixed treats	Alex says “eight”
33	7 1” Blocks	Alex says “sih–none”

^a “Mixed” stimuli were of varying sizes within sets, randomized across sets. Only the 1” blocks and the small beans were sets that consisted of a single sized stimulus.

session/wk with human models, to learn to order the Arabic numerals as “6 < 7 < 8”. With no further training, he then integrated these numerals’ ordinality with all those he knew (“1–6”)—the representation of serial order necessary for the induction that was the focus of this study.

We could then ask the comparative question: Could a nonhuman infer the cardinal value of new numerals (i.e., “sih–none”, “eight”) from their ordinal relations to other numerals whose cardinal values were already known? The answer was “yes”. In comprehension, Alex nearly perfectly indicated 7-item sets if asked for “seven” and 8-item sets if asked for “eight”, without any training to associate those numerals with specific set sizes. In production, he was well above chance in labeling sets of 7 as “sih–none” and 8 as “eight”, again without any training to enumerate 7- or 8-item sets. Note, he could not simply have associated a new utterance with a novel numerical set, because he had simultaneously learned two new number labels.

As noted above, previous studies had shown nonhumans capable of representing arbitrary ordered lists, bringing two distinct ordered lists into correspondence, being sensitive to ordinal relations among quantities represented as analog magnitudes or with parallel individuation, and mapping numerals to non-linguistic representations of number. But no other nonhuman has made Alex’s

Table 6

Comprehension data for collections of objects; questions are “What color X?”.

Collection ^a	Question	Response
4B, 7Y, 10Gy blocks	What color 7?	Yellow
20, 4G, 7Ppl sticks	What color 4?	Green
	What color 7?	Purple
5Y, 8R, 10G paper pieces	What color 8?	Rose
3B, 7O, 8G wool pompons	What color 7?	Orange
6Y, 7G, 8B candies	What color 8?	Six-yellow, blue
6G, 7O, 8Y jelly beans	What color 7?	Yellow
7B, 8Gy, 9O wool pompons	What color 8?	Eight (4×), grey
6G, 7O, 9R pasta	What color 7?	Orange
5Ppl, 6O, 7G jelly beans	What color 7?	Green
4Gy, 7O, 8B blocks	What color 8?	Blue
4B, 7Y, 10Ppl wool	What color 7?	Yellow
7R, 8G, 9Ppl wool	What color 8?	Green

^a B refers to blue, G to green, Gy to grey (all shades of grey to black), O to orange, Ppl to purple, R to rose (red), and Y to yellow. Again, except for blocks, items are all differently-sized.

induction. As apes learned to correlate specific Arabic symbols with exact sets of items (Ai 0–9, Sheba, 0–8: Biro & Matsuzawa, 2001; Boysen, 1993; Boysen & Berntson, 1989; Matsuzawa, 1985; Matsuzawa, 2009; Murofushi, 1997), no savings in training trials occurred for larger numerals; separate training was needed to order numerals

(Boysen et al., 1993; Kawai & Matsuzawa, 2000; Matsuzawa et al., 1991; Tomonaga & Matsuzawa, 2000; Tomonaga, Matsuzawa, & Itakura, 1993). However, neither of these apes had the occasion to induce how the list represents number, because neither was taught an ordered list before being taught cardinal values of each numeral. Alex's logical and symbolic capacities may be widely shared by nonhumans with appropriate training.

The present study provides the first demonstration that a nonhuman can engage in the bootstrapping process that underlies the construction of the integer list representation of number, but it is the first study in which an animal was given the chance to display this ability. Matsuzawa, however, could test whether knowing the integer list prior to being taught cardinal meanings of integers is sufficient to support the induction Alex made. Several of his young chimpanzees order numerals from “1”–“9” (Matsuzawa, 2009) but, unlike Ai or Sheba, have not yet been taught set sizes that correspond to the numerals in their list. Such training is clearly possible. If chimps first taught the ordered list of numerals can make the same induction as can Alex and young children, at some point in learning this sequence they should figure out cardinal meanings of higher numerals from their place in the list. Unlike Ai and Sheba, they should need far fewer trials to learn cardinal values of numerals greater than “4” than for “1”–“4”. Alternatively, it may be necessary to both produce and comprehend mappings between small numerals and set sizes (a step that took extra training for Ai and Sheba, and had already been achieved by Alex) before being taught cardinal meanings of larger numerals for any savings to be observed. Such research would bear on whether the capacity to induce cardinal meanings of numerals from their position in a count list is widespread among nonhumans, and would bear on what prior experience supports this inductive process.

5.1. Alex's induction

Alex's induction was not fully warranted by his training. True, learning the Arabic numeral ordering “ $6 < 7 < 8$ ” warranted the inference that “7” and “8” are numerals denoting particular cardinal values, as is “6”; 7 and 8 would thus be expected to participate in previously learned explicit relations among numbers. But these relations are consistent with Arabic numerals (“7, 8”) and their vocal labels (“sih–none, eight”) denoting any ordered cardinal values >6 (e.g., 10 or 20, respectively), as long as the value of “7/sih–none” was less than that of “8/eight”. Thus, no deductively valid inference exists from what he had been taught to the cardinal values denoted by “7/sih–none” and “8/eight” being 7 and 8, respectively. Nonetheless, this is the inductive inference Alex made.

We must compare Alex's inductive inferences to those of children when they induce how the count list works (Rips, Asmuth, & Bloomfield, 2008; Rips, Bloomfield, & Asmuth, 2008). To infer what he did, Alex must have generalized the relations between the first six ordered numerals (i.e., that each denote a set one greater than the immediately preceding numeral) to the newly taught ones—that these next new numerals denoted the next cardinal values in the series.

Alex's previous training might have assisted this inductive process. Numerals had always been taught in pairs referring to sets differing from each other and/or known numerals by exactly 1. That is, he first learned set sizes to which “three” and “four” referred, then was taught “five” ($4 + 1$) and “two” ($3 - 1$). After mastering these labels, he learned “six” ($5 + 1$) and “one” ($2 - 1$). He also inferred that “none” was less than “one” (Pepperberg & Gordon, 2005). Too, he had been extensively tested on his knowledge of the numerical ordering of “1/one” through “6/sih”, where adjacent pairs in the ordered list differ by one. This experience may have led him to restrict his hypotheses about the numerical relations between “6, 7”, and “8” to set sizes that differ from each other by just 1.

However his inductive inference was appropriately constrained, his inference reflects an implicit representation of the successor function, and requires appreciation of relations between two different “follows” relations—next in an ordered series and next in a series of sets related as n and $n + 1$. These are among the logical foundations of integer concepts; Alex's deployment of them undermines claims of human uniqueness with respect to these computational resources.

5.2. The format of Alex's number representations

Alex induced cardinal values of “7/sih–none” and “8/eight” from his knowledge of cardinal meanings of “1/one” through “6/sih–sih” plus that of ordinal relations among the numerals “1”–“8”. The data leave open which prelinguistic numerical representations supported this induction—analogue magnitude representations or the machinery of parallel individuation and computations of 1–1 correspondence. Possibly, for Alex, analogue magnitude representations underlie the meaning of “1”–“6”, and mappings between the order in the list and the order of cardinals were supported by the ordinal relations among analogue magnitude values. Alex's representations of the cardinal meanings of “6/sih–sih”, “7/sih–none”, and “8/eight” were precise enough for him to distinguish each of these cardinal values from each other. Furthermore, in the comprehension test (Table 6), half the trials included sets of 9 or 10, as well as sets of 7 and 8. When probed “What color seven?” or “What color eight?”, Alex never once gave the color of the 9-item or 10-item set, despite the fact that these sets are approximately 8, and satisfy the learned relation of being >6 . Thus, if the representations underlying the cardinal meanings of these numerals are analogue magnitude representations, they are very sharply tuned. Clearly, analogue magnitude representations can support ordinal judgments (e.g., Brannon & Terrace, 1998, 2000; Dehaene, 2009; Gallistel & Gelman, 1992, 2005), and although we do not know the precision of set comparisons Alex's analogue magnitude representations allow, human adults can often resolve differences in the ratio of 8:9 (Barth et al., 2003; Halberda & Feigenson, 2008).

Nonetheless, two considerations lead us to favor the hypothesis that some representations based on 1–1 correspondence and “exactly N” (rather than “approximately N”) supported the meanings Alex assigned to numerals.

First, and most important, analog magnitude representations do not naturally support the successor function (Carey, 2001, 2009; Leslie, Gelman, & Gallistel, 2008; Rips, Asmuth, et al., 2008), but Alex's induction was that "sih-sih", "sih-none", and "eight" are related as are "two", "three", and "four". Each successive numeral labels sets with cardinal values one greater than its predecessor. Analog magnitude representations fail to support this induction for two reasons: First, they contain no representation of "exactly one". Second, they are naturally compared by computing ratios, not intervals. The difference between 2 and 3 is experienced as the same as that between 6 and 9, not as that between 6 and 7. Thus, while analog magnitude representations support addition and subtraction, they do not make salient the induction Alex actually made.

Second, one aspect of our data bears dwelling upon. In 13 trials of the Cardinal Induction Production Test (Table 5) 1 or 2 items were added or subtracted from a set he had already labeled. On all these trials, Alex made only one error at reporting the numeral that referred to the transformed set (Trial 3), which he immediately self-corrected. Alex's success, which was higher than his success rate at production overall, suggests he represented the relation between successive numerals as +1. Again, while analog magnitude representation support addition and subtraction (including by 1 and 2), these data lend support to the additional conclusion that Alex's numerals were related to each other via the successor function, which, as argued above, does not naturally fall out of analog magnitude representations.

We acknowledge that these arguments do not conclusively show that the nonhumans extensively trained to map external numerals to cardinal values of sets created representations that go beyond their analog magnitude system by drawing on representations that implement the successor function based on 1–1 correspondence. The important result from this study is that Alex created cardinal representations of "7/sih-none" and "8/eight" from their position in the list of ordered numerals together with his knowledge of the meanings of "1/one" through "6/sih".

5.3. Similarity to and differences from children's acquisition of numeral representations

Alex's induction is what separates "subset-knowers" from "CP knowers". Subset-knowers have assigned cardinal meaning to only a subset of numerals in their count-list. CP knowers have previously worked out cardinalities to which "one" through "four" refer, have induced how counting works, and can assign cardinal meanings to any numeral in their count list (Carey, 2009).

Nonetheless, Alex's route to mastering a numeral list representation of integers differed from that of children. He was never taught to recite a count list; he induced relations of numerical order among his numerals after learning cardinal meanings of "one" through "six" as well as the labels for Arabic numerals. Children learn to recite the count list, and engage in a count routine, even before becoming "one-knowers". Alex did not learn numerals in exact order; he did not go through the knower-levels. He was taught that "four" represented exactly 4 items, and no data exist

that he ever took "four" to refer to any/all cardinalities above 3, the meaning assigned to "four" by a child "three"-knower. Moreover, children's assigning numerical meaning to the first numerals in the count list is aided by numerals' status as quantifiers (e.g., Barner, Libenson, Cheung, & Takasaki, 2009; Bloom & Wynn, 1997; Sarnecka, Kamenskaya, Ogura, Yamana, & Yudovina, 2007), information unavailable to Alex. Clearly, many routes could lead to creating a numeral list representation of integers, but both Alex and young children clearly can master this culturally constructed system. What parrots—and we venture any other nonhuman—would seem to lack is the capacity to invent it in the first place.

Acknowledgements

We thank Elizabeth Spelke for her valuable comments and suggestions. This study was supported in part by The Radcliffe Institute for Advanced Study and written with support from NSF grant BCS-0920878 (I.M.P. under Ken Nakayama) and NIH grant 2R01 HD038338 (S.C.); further support came from donors to *The Alex Foundation* (particularly the Pearl Family Foundation, Anita Keefe, the Marc Haas Foundation, Janice Boyd, Anders and Joyce Sterner, Joseph Golden, Alex and Michael Shuman, Pat Hill, William Lipton, Nancy Chambers, Greg LaMorte, Kathryn and Walter McAdams, Megumi Oka and the Makioka Foundation, Elva and Bob Mathieson, Jay Sugarman, Debby and Michael Smith, Rose and Ken Lawlor, George and Lisa Bono, Regina and Ed Underwood, John Paton, Deborah Rivel and Robert Goodale, Ann Arbor Cage Bird Club, Mark Hagen, the Canadian Parrot Society, Debbie Rijnders/Tinley Advies & Producties BV, Vanessa Rolfe, LeighAnn and Carl Hartsfield, Mary Ogg, Featherlust Farms Bird Shop, Reno Area Avian Enthusiasts, Janet and Kip Trumbule, Bird Paradise, the Medwick Foundation, Tammy Carreiro, Carol Samuelson, Nancy Clark and Bill Broach, Rocky Mountain Society of Aviculture, Friends of Greys, the Andrew De Mar Family Foundation, Tracy Hylka, Charlene Bebko, Ellen Paul, Hunt Associates, Mary Long). We thank the Harrison family for Harrison's Bird Diet, Fowl Play for treats, Carol D'Arezzo for Alex's perch, and Avian Adventures for cages and additional financial assistance. Betsy Lindsay, Jen Vicinay, Jessica Balter, Jessica Vitale, Chelsea Canan, Vincent Capanelli assisted with testing.

References

- Barner, D., Libenson, A., Cheung, P., & Takasaki, M. (2009). Cross-linguistic relations between quantifiers and numerals in language acquisition: Evidence from Japanese. *Journal of Experimental Child Psychology*, *103*, 421–440.
- Barth, H., Kanwisher, N., & Spelke, E. (2003). The construction of large number representations in adults. *Cognition*, *86*, 201–221.
- Beran, M. J., Harris, E. H., Evans, T. A., Klein, E. D., Chan, B., Flemming, T. M., et al. (2008). Ordinal judgements of symbolic stimuli by capuchin monkeys (*Cebus paella*) and rhesus monkeys (*Macaca mulatta*): The effects of differential and nondifferential reward. *Journal of Comparative Psychology*, *122*, 52–61.
- Biro, D., & Matsuzawa, T. (2001). Use of numerical symbols by the chimpanzee (*Pan troglodytes*): Cardinals, ordinals, and the introduction of zero. *Animal Cognition*, *4*, 193–199.
- Bloom, P., & Wynn, K. (1997). Linguistic cues in the acquisition of number words. *Journal of Child Language*, *24*(3), 511–533.

- Bowmaker, J. K., Heath, L. A., Das, D., & Hunt, D. M. (1994). Spectral sensitivity and opsin structure of avian rod and cone visual pigments. *Investigative Ophthalmology and Visual Science*, 35, 1708.
- Bowmaker, J. K., Heath, L. A., Wilkie, S. E., Das, D., & Hunt, D. M. (1996). Middle-wave cone and rod visual pigments in birds: Spectral sensitivity and opsin structure. *Investigative Ophthalmology & Visual Science*, 37, S804.
- Boysen, S. T. (1993). Counting in chimpanzees: Nonhuman principles and emergent properties of number. In S. T. Boysen & E. J. Capaldi (Eds.), *The development of numerical competence: animal and human models* (pp. 39–59). Hillsdale, NJ: Erlbaum.
- Boysen, S. T., & Berntson, G. G. (1989). Numerical competence in a chimpanzee (*Pan troglodytes*). *Journal of Comparative Psychology*, 103, 23–31.
- Boysen, S. T., Berntson, G. G., Shreyer, T. A., & Quigley, K. S. (1993). Processing of ordinality and transitivity by chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, 107, 208–215.
- Brannon, E. M., & Terrace, H. S. (1998). Ordering of the numerosities 1–9 by monkeys. *Science*, 282, 746–749.
- Brannon, E. M., & Terrace, H. S. (2000). Representation of the numerosities 1–9 by rhesus macaques. *Journal of Experimental Psychology: Animal Behavior Processes*, 26, 31–49.
- Carey, S. (2001). Cognitive foundations of arithmetic: Evolution and ontogenesis. *Mind and Language*, 16, 37–55.
- Carey, S. (2009). *The origin of concepts*. New York: Oxford.
- Chen, S., Swartz, K., & Terrace, H. S. (1997). Knowledge of the ordinal position of list items in rhesus monkeys. *Psychological Science*, 8, 80–86.
- D'Amato, M. R., & Colombo, M. (1988). Representation of serial order in monkeys (*Cebus paella*). *Journal of Experimental Psychology: Animal Behavior Processes*, 14, 131–139.
- Dehaene, S. (2009). Origins of mathematical intuitions: The case of arithmetic. *Annals of the New York Academy of Sciences*, 1156, 232–259.
- Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Science*, 8, 307–314.
- Frank, M. C., Everrett, D. L., Fedorenko, E., & Gibson, E. (2008). Number as a cognitive technology: Evidence from Pirahã language and cognition. *Cognition*, 108, 819–824.
- Fuson, K. (1988). *Children's counting and concepts of number*. New York: Springer-Verlag.
- Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition*, 44, 43–74.
- Gallistel, C. R., & Gelman, R. (2005). Mathematical cognition. In K. S. Holyoak, Jr. & R. G. Morrison, Jr. (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 559–588). Cambridge: Cambridge University Press.
- Gordon, P. (2004). Numerical cognition without words: Evidence from Amazonia. *Science*, 306, 496–499.
- Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the “number sense”: The approximate number system in 3-, 4-, 5-, and 6-year-olds and adults. *Developmental Psychology*, 44, 1457–1465.
- Harris, E. H., Beran, M. J., & Washburn, D. A. (2007). Ordinal-list integration for symbolic, arbitrary, and analog stimuli by rhesus macaques (*Macaca mulatta*). *Journal of General Psychology*, 134, 183–197.
- Harris, E. H., Gulledge, J. P., Beran, M. J., & Washburn, D. A. (2010). What do Arabic numerals mean to macaques (*Macaca mulatta*)? *Journal of Experimental Psychology: Animal Behavior Processes*, 36, 66–76.
- Hurford, J. R. (1987). *Language and number*. Cambridge: Cambridge University Press.
- Kawai, N., & Matsuzawa, T. (2000). Numerical memory span in a chimpanzee. *Nature*, 403, 39–40.
- Klahr, D., & Wallace, J. G. (1973). The role of quantification operators in the development of conservation of quantity. *Cognitive Psychology*, 4, 301–327.
- Le Corre, M., Van de Walle, G., Brannon, E. M., & Carey, S. (2006). Revisiting the competence/performance debate in the acquisition of the counting principles. *Cognitive Psychology*, 52, 130–169.
- Leslie, A. M., Gelman, R., & Gallistel, C. R. (2008). The generative basis of natural number concepts. *Trends in Cognitive Science*, 12, 213–218.
- Matsuzawa, T. (1985). Use of numbers by a chimpanzee. *Nature*, 315, 57–59.
- Matsuzawa, T. (2009). Symbolic representation of number in chimpanzees. *Current Opinions in Neurobiology*, 19, 92–98.
- Matsuzawa, T., Itakura, S., & Tomonaga, M. (1991). Use of numbers by a chimpanzee: A further study. In A. Ehara, T. Kimura, O. Takenaka, & M. Iwamoto (Eds.), *Primate today* (pp. 317–320). Amsterdam: Elsevier.
- Murofushi, K. (1997). Numerical matching behavior by a chimpanzee (*Pan troglodytes*): Subitizing and analogue magnitude estimation. *Japanese Psychological Research*, 39, 140–153.
- Patterson, D. K., & Pepperberg, I. M. (1994). A comparative study of human and parrot phonation: I. Acoustic and articulatory correlates of vowels. *Journal of the Acoustical Society of America*, 96, 634–648.
- Patterson, D. K., & Pepperberg, I. M. (1998). A comparative study of human and Grey parrot phonation: II. Acoustic and articulatory correlates of stop consonants. *Journal of the Acoustical Society of America*, 103, 2197–2213.
- Pepperberg, I. M. (1981). Functional vocalizations by an African Grey parrot. *Zeitschrift für Tierpsychologie*, 55, 139–160.
- Pepperberg, I. M. (1983). Cognition in the African Grey parrot: Preliminary evidence for auditory/vocal comprehension of the class concept. *Animal Learning & Behavior*, 11, 179–185.
- Pepperberg, I. M. (1987). Evidence for conceptual quantitative abilities in the African Grey parrot: Labeling of cardinal sets. *Ethology*, 75, 37–61.
- Pepperberg, I. M. (1990a). Cognition in an African Grey parrot (*Psittacus erithacus*): Further evidence for comprehension of categories and labels. *Journal of Comparative Psychology*, 104, 41–52.
- Pepperberg, I. M. (1990b). Referential mapping: Attaching functional significance to the innovative utterances of an African Grey parrot (*Psittacus erithacus*). *Applied Psycholinguistics*, 11, 23–44.
- Pepperberg, I. M. (1992). Proficient performance of a conjunctive, recursive task by an African Grey parrot (*Psittacus erithacus*). *Journal of Comparative Psychology*, 106, 295–305.
- Pepperberg, I. M. (1994). Evidence for numerical competence in an African Grey parrot (*Psittacus erithacus*). *Journal of Comparative Psychology*, 108, 36–44.
- Pepperberg, I. M. (1999). *The Alex studies: Cognitive and communicative abilities of Grey parrots*. Cambridge, MA: Harvard University Press.
- Pepperberg, I. M. (2006a). Addition by a Grey Parrot (*Psittacus erithacus*), including absence of quantity. *Journal of Comparative Psychology*, 120, 1–11.
- Pepperberg, I. M. (2006b). Ordinality and inferential abilities of a Grey Parrot (*Psittacus erithacus*). *Journal of Comparative Psychology*, 120, 205–216.
- Pepperberg, I. M. (2007). Grey parrots do not always ‘parrot’: Phonological awareness and the creation of new labels from existing vocalizations. *Language Sciences*, 29, 1–13.
- Pepperberg, I. M., & Brezinsky, M. V. (1991). Acquisition of a relative class concept by an African Grey Parrot (*Psittacus erithacus*): Discriminations based on relative size. *Journal of Comparative Psychology*, 105, 286–294.
- Pepperberg, I. M., & Gordon, J. D. (2005). Number comprehension by a Grey Parrot (*Psittacus erithacus*), including a zero-like concept. *Journal of Comparative Psychology*, 119, 197–209.
- Pepperberg, I. M., & Kozak, F. A. (1986). Object permanence in the African Grey parrot (*Psittacus erithacus*). *Animal Learning & Behavior*, 14, 322–330.
- Pepperberg, I. M., & Lynn, S. K. (2000). Perceptual consciousness in Grey parrots. *American Zoologist*, 40, 393–401.
- Pepperberg, I. M., Vicinay, J., & Cavanagh, P. (2008). The Müller-Lyer illusion is processed by a Grey Parrot (*Psittacus erithacus*). *Perception*, 37, 765–781.
- Pepperberg, I. M., & Wilkes, S. R. (2004). Lack of referential vocal learning from LCD video by Grey parrots (*Psittacus erithacus*). *Interaction Studies*, 5, 75–97.
- Pica, P., & Lecompte, A. (2008). Theoretical implications of the study of numbers and numerals in Mundurucu. *Philosophical Psychology*, 21, 507–522.
- Premack, D. (1976). *Intelligence in ape and man*. Hillsdale, NJ: Erlbaum.
- Rips, L. J., Asmuth, J., & Bloomfield, A. (2008). Do children learn the integers by induction? *Cognition*, 106, 940–951.
- Rips, L. J., Bloomfield, A., & Asmuth, J. (2008). From numerical concepts to concepts of number. *Behavioral and Brain Sciences*, 31, 623–687.
- Sarnecka, B. W., & Carey, S. (2008). How counting represents number: What children must learn and when they learn it. *Cognition*, 108, 662–674.
- Sarnecka, B. W., Kamenskaya, V. G., Ogura, T., Yamana, Y., & Yudovina, J. B. (2007). From grammatical number to exact numbers: Early meanings of “one”, “two”, and “three”, in English, Russian, and Japanese. *Cognitive Psychology*, 55(2), 136–168.
- Scarf, D., Hayne, H., & Colombo, M. (2011). Pigeons on par with primates in numerical competence. *Science*, 334, 1664.
- Spaepen, E., Coppola, M., Spelke, E. S., Carey, S. E., & Goldin-Meadow, S. (2011). Number without a language model. *Proceedings of the National Academy of Sciences*, 108, 3163–3168.

- Swartz, K. B., Chen, S., & Terrace, H. S. (2000). Serial learning by rhesus monkeys: II. Learning four-item lists by trial and error. *Journal of Experimental Psychology: Animal Behavior Processes*, 26, 274–285.
- Terrence, H. S., Son, L. K., & Brannon, E. M. (2003). Serial expertise of rhesus macaques. *Psychological Science*, 14, 66–73.
- Tomonaga, M., & Matsuzawa, T. (2000). Sequential responding to Arabic numerals with wild cards by the chimpanzee (*Pan troglodytes*). *Animal Cognition*, 3, 1–11.
- Tomonaga, M., Matsuzawa, T., & Itakura, S. (1993). Teaching ordinals to a cardinal trained chimpanzee. *Primate Research*, 9, 67–77.
- Wynn, K. (1990). Children's understanding of counting. *Cognition*, 36, 155–193.
- Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cognitive Psychology*, 24, 220–251.