Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

HHS Public Access
Author manuscript
J Comp Psychol. Author manuscript; available in PMC 2016 December 12.

Published in final edited form as:
J Comp Psychol. 2010 November ; 124(4): 369–380. doi:10.1037/a0019855.

Can nonhuman primates use tokens to represent and sum quantities?

Theodore A. Evans,
Language Research Center, Georgia State University

Michael J. Beran, and
Language Research Center, Georgia State University

Elsa Addessi
Unit of Cognitive Primatology and Primate Center, Institute of Cognitive Sciences and Technologies, CNR

Abstract

It is unclear whether nonhuman animals can use physical tokens to flexibly represent various quantities by combining token values. Previous studies showed that chimpanzees (*Pan troglodytes*) and a macaque (*Macaca mulatta*) were only partly successful in tests involving sets of different-looking food containers representing different food quantities, while some capuchin monkeys (*Cebus apella*) have shown greater success in tests involving sets of various concrete objects representing different food quantities. Some of the discrepancy in results between these studies may be attributed to the different methods employed. In an effort to reconcile these discrepancies, we presented two primates species, chimpanzees and capuchin monkeys, with two token tasks. The critical test in each task involved summing the value of multiple tokens of different types to make accurate quantity judgments. We found that, using either method, individuals of both species learned to associate individual tokens with specific quantities, as well as successfully compare individual tokens to one another or to sets of visible food items. However, regardless of method, only a few individuals exhibited the capacity to sum multiple tokens of different types and then use those summed values to make an optimal response. This suggests that flexible combination of symbolic stimuli in quantity judgments tasks is within the abilities of chimpanzees and capuchins but does not characterize the majority of individuals. Furthermore, the results suggest the need to carefully examine specific methodological details that may promote or hinder such possible representation.

Humans learn very early in life to use symbols to represent objects, actions, and other relevant types of information. Symbol use confers many advantages for our species, notably in the ways that it affords complex decision-making and the representation of past and future events (e.g., Deacon, 1997). Although once considered a uniquely human capacity, it is now recognized that some nonhuman species can use symbols in ways analogous to some forms

Address correspondence to the first author at Language Research Center, Georgia State University, P.O. Box 5010, Atlanta, GA 30302-5010., theodore.evans@gmail.com.
Theodore A. Evans, Language Research Center, Georgia State University; Michael J. Beran, Language Research Center, Georgia State University; Elsa Addessi, Unit of Cognitive Primatology and Primate Center, Institute of Cognitive Sciences and Technologies, CNR.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

of human symbol use. Some animals have been taught to associate arbitrary symbols with real world referents. The most dramatic examples of symbol use by nonhuman animals come from projects designed to investigate language acquisition (e.g., Gardner & Gardner, 1969; Herman, Richards, & Wolz, 1984; Pepperberg, 1999; Premack & Premack, 1983; Rumbaugh, 1977; Rumbaugh & Washburn, 2003; Savage-Rumbaugh, 1986; Shusterman & Gisiner, 1988). Evidence from these projects indicates that bonobos, chimpanzees, parrots, dolphins and sea lions can flexibly use symbols to accomplish a variety of goals.

Of particular interest is the ability of nonhuman animals to use symbols to represent quantities and in some cases to approximate the counting routines used by humans. Here, too, a variety of species have shown numerical symbol competency, including chimpanzees (e.g., Beran & Rumbaugh, 2001; Boysen & Berntson, 1989; Boysen & Hallberg, 2000; Matsuzawa, 1985; Tomonaga & Matsuzawa, 2000), a parrot (e.g., Pepperberg, 1987, 1994, 2006), dolphins (Mitchell, Yao, Sherman, & O'Regan, 1985), and various monkeys species (e.g., Olthof, Iden, & Roberts, 1997; Olthof & Roberts, 2000; Washburn & Rumbaugh, 1991).

Other evidence of symbolic-like capacities in nonhuman species comes from studies that use tokens. For example, capuchin monkeys (hereafter capuchins) have been trained to associate different types of tokens with different food items (Addessi, Crescimbene, & Visalberghi, 2007, 2008a; Brosnan & de Waal, 2004) and different tool objects (Westergaard, Liv, Chavanne, & Suomi, 1998), and to use such tokens in exchange tasks with other monkeys (Westergaard, Evans, & Howell, 2007). Capuchins and chimpanzees have been presented with tasks that assess reactions to social inequity (Brosnan & de Waal, 2003; Brosnan, Schiff, & de Waal, 2005), preference transitivity (Addessi, Mancini, Crescimbene, Padoa-Schioppa, & Visalberghi, 2008b), loss aversion (Chen, Lakshminaryanan, & Santos, 2006), and endowment effects (Brosnan et al., 2007, Lakshminaryanan, Chen, & Santos, 2008). Most evidence suggests that these species do learn to, at minimum, associate tokens with certain foods or certain task outcomes, and even to treat tokens as symbols that stand for those foods or outcomes (Addessi et al., 2008b; Brosnan & Beran, 2009).

Beran, Beran, Harris, and Washburn (2005) presented chimpanzees and a rhesus macaque (hereafter macaque) with a different variation of a quantity judgment task with symbols. These animals first chose between pairs of colored plastic eggs where each color egg always contained a specific and unique number of items (pink always contained five items, blue always contained four items, green always contained three items, orange always contained two items, and yellow always contained one item). All animals rapidly learned to select the egg in each pairing that contained more food. This result matched that of other studies using arbitrary stimuli that were associated with specific food items, including tests of real world objects and computer stimuli. For example, dolphins performed well in choosing the more valuable (i.e., larger quantity) item when given pairs to choose from (Mitchell et al., 1985), and macaques learned to choose the larger of two Arabic numerals to maximize food reward (Washburn & Rumbaugh, 1991). Also, chimpanzees have selected Arabic numerals of a lower value in order to solve a reverse-reward contingency task (Boysen, Mukobi & Bernston, 1999), and a grey parrot has verbalized which of two presented Arabic numerals is "bigger" (Pepperberg, 2006).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

The chimpanzees in the Beran et al. (2005) study also performed at high levels when comparing an egg to a visible set of food items (e.g., choosing a blue egg over three visible items but not five visible items; choosing a green egg over two visible items but not four visible items). They were successful with comparisons in which they had to consider the number of eggs presented in comparison to visible items (e.g., choosing two blue eggs with eight items total over six visible items, but choosing six visible items over one blue egg). However, all subjects failed in one critical test: comparing two eggs of a smaller individual value but larger total value to one egg of a larger individual value but a smaller total. For example, when presented with two blue eggs (that would contain eight total items) and a pink egg (with five items), the animals chose the pink egg consistently. Performance on these kinds of trials was significantly below chance, and various attempts to facilitate higher performance also failed (e.g., putting the eggs into opaque containers). Thus, when comparing these quantity symbols to visible quantities, all animals performed very well, and they performed very well when comparing any two eggs to each other, but they could not accommodate two pieces of information (the egg color and the number of eggs of that color) when all presented items were these symbols.

This was a striking failure on the part of the chimpanzees and macaque. Beran et al. (2005) proposed that the animals experienced a Stroop-like effect wherein they attended only to the more salient cue (egg color) and not the less salient cue (number of eggs of that color). At minimum, they struggled presumably because of response competition between physical egg quantity and quantity of items represented by each egg. However, they may also have struggled because this condition placed the largest load on memory, with the animals having to remember the quantities of each egg type as well as the summed amount from the multiple eggs that could be presented in one color.

In contrast to the results from Beran et al. (2005), Addessi et al. (2007) reported that capuchins could sum and compare sets of tokens that represented different quantities. In their task, capuchins first learned to distinguish one type of token that always gave the animals three food items, when exchanged, from another type of token that always gave them one item. Capuchins then were presented with one of the 3-item tokens and one to five single-item tokens in the alternate choice set. Four of ten capuchins performed well in choosing the larger overall amount of food. Another four capuchins preferred the three-item token in all comparisons with single-item tokens, a result that we view as comparable to the failure of chimpanzees and a macaque in the Beran et al. (2005) condition in which egg color cues were in conflict with total food amount. The remaining two capuchins in the Addessi et al. (2007) experiment preferred the 3-item token only when presented in comparison to one single-item token (the training condition), whereas in all other conditions these two monkeys preferred the set that had the greater number of tokens.

Importantly, Addessi et al. (2007) reported that two capuchins also succeeded on a subsequent test in which they faced a binary choice between one or two of the three-item tokens and three to six single-item tokens. This suggested that capuchins estimated the value of individual token quantities and then combined the representation of quantities before making a choice.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Thus, there remains ambiguity regarding whether nonhuman primates can take into account both the individual value of tokens and the quantity of those tokens presented when performing quantity judgment tasks. The data indicate that capuchins can perform this task, but chimpanzees and macaques cannot. However, it is important to note that only some capuchins were successful in the Addessi et al. (2007) experiment, and only a single macaque and four chimpanzees were tested in the Beran et al. (2005) experiment. In addition, there were substantial methodological differences between these studies that might account for the different outcomes. Beran et al. (2005) used five different token values whereas Addessi et al. (2007) used only two. Addessi et al. (2007) presented more than two tokens of the lesser value in some of their conditions, whereas Beran et al. (2005) did not (except in their final experiment where one of the chimpanzees improved albeit to only chance levels from below chance levels). Thus, a more direct comparative assessment is required to discern to what extent nonhuman animals can perform the symbol summation task.

We presented two of the species in question, chimpanzees and capuchins, with both variations of established symbol summation tasks. The chimpanzees already had completed the Beran et al. (2005) task and so were only given tests similar to those administered by Addessi et al. (2007). Capuchins from the same laboratory (the Language Research Center) were presented with the Beran et al. (2005) tests, as well as the Addessi et al. (2007) tests. These new data would help assess the extent to which a specific method may lead to either success or failure in symbol summation and what differences in the tasks may support or prevent successful performance.

Experiment 1: Can capuchins perform ordinal and summative judgments of containers holding different quantities of discrete food items?

We closely followed the methods of the original study conducted by Beran et al. (2005) so that we would have an accurate comparison of species on this test. The only differences in methodology between the original study and the current one were the type of food items used and the features used to differentiate food containers. These methodological differences reflected differences between species' caloric requirements and perceptual capacities, respectively. Despite these small changes, we hypothesized that capuchins would perform similarly to the previously tested chimpanzees and rhesus macaque, given their similar performance in other studies of a quantitative nature (e.g., Beran, Evans, Leighty, Harris, & Rice, 2008a; Judge, Evans, & Vyas, 2005).

Methods

Participants—We observed 4 capuchin monkeys (*Cebus apella*): Liam (male, 3 years old), Wren (female, 4 years old), Nala (female, 4 years old), and Lily (female, 10 years old). All monkeys participated in one or more previous studies involving quantity judgments. In two of these studies, they ordered completely visible arrays of 1 to 10 digital objects (Beran, 2008) or completely visible arrays of digital numerals within the range of 0 to 9 (Beran et al., 2008b). In two other studies, these monkeys chose between two sets of 1 to 5 food items that were visible either one whole set at a time (Beran et al., 2008a) or one item at a time

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

(Evans, Beran, Harris & Rice, 2009) prior to selection. Thus, prior to the present study, these monkeys had never been evaluated for the ability to choose between individual containers or between multiple containers based on pre-trial experience with the different containers' contents.

Design and Procedure—We separated participants into individual enclosures for testing. Monkeys observed the presentation of choice options through a clear plastic panel that was attached to the side of the enclosure facing the experimenters. The panel had two holes large enough for the monkeys to reach through in order to respond to stimuli. These holes forced the monkeys to choose only one of the two choice options, and the monkeys all had previous experience learning that they could only reach through one hole at a time, or a response would not be rewarded (e.g., Beran et al., 2008a). All stimuli were presented on the top of a utility cart that could be rolled in and out of the monkeys' reach. The stimuli consisted of white wooden boxes with hinged lids. There were five types of boxes, each with a unique shape and sticker pattern. Each box type contained a different quantity (1, 2, 3, 4 or 5) of 45-mg banana-flavored pellets (Bio-Serv, Frenchtown, NJ). We randomly assigned quantities to box types (shapes/patterns) at the beginning of the study and the monkeys had no experience with these box types prior to this study.

Two experimenters participated in test sessions. At the beginning of a trial, Experimenter 1 (the first author) positioned the cart out of reach of the monkey and displayed two choice options (some combination of pre-baited containers and/or visible food items, depending on the condition). The two choices always were presented on opposite ends of the cart, each in a large bowl directly in line with one of the holes in the clear panel (thus, the monkeys never saw the food items directly in any trials as they were already in the containers). Experimenter 1 presented the choices to the monkey by first closing his eyes and then slowly rolling the cart forward until it touched the test enclosure. He did so using only two fingers of one hand, pressed against the center of the cart shelf, and thus could only control the forward motion of the cart (if the cart swayed or angled slightly in favor of one response option, this was completely random and out of control of Experimenter 1). Experimenter 2 (the second author) stood 2–3 m to the side of the test enclosure, where the monkey could not see him, and he could only see the monkey's hands reach out of the holes in the panel. From that position, Experimenter 2 called out the first choice option touched by the monkey. Experimenter 1 then rolled the cart back, opened the selected container(s) and poured its contents into the bowl that previously held the container(s). Experimenter 1 then allowed the monkey to consume the food items in the selected bowl while Experimenter 2 re-baited the containers and stowed them out of view of the monkey. In this and all subsequent experiments, we randomized which end of the cart held the choice option with the greater amount of food, as well as exactly what comparison was presented on a given trial.

Phase 1: One Container vs. One Container: In this phase the monkeys received comparisons of two individual containers. We presented all combinations of the five different container types (a total of 10 unique comparisons). Each session consisted of 40 trials in which we presented an equal number of each comparison. We presented each monkey with one session per day, three days per week, until each monkey reached an 85%

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

selection criterion of the correct container (i.e., the container holding more food items) within a single test session.

Phase 2: One or Two Container(s) vs. Two Containers: In this phase we presented monkeys with the choice between 1 or 2 containers of one type and 2 containers of another type. Thus, the number of food items presented in each choice option ranged from 1 to 10. We presented 6 trial types that differed on the basis of the availability of two stimulus cues, number of containers and container value (i.e., contents), and on the basis of the relevance of those cues to the relative amount of food contained in the two choice options (see also Table 1). In *Concordant* trials, container number and container value both differed between the choice options, and both cues were consistent with the relative amount of food in the choice options. In *Number* trials, only container number differed between the choice options and was consistent with the relative amount of food in each option. In *Value* trials, only container value differed between the choice options and was consistent with the relative amount of food in each option. In *Discordant – Number* trials, container number and container value both differed between the choice options, but only container number was consistent with the relative amount of food in the choice options. In *Discordant – Value* trials, container number and container value both differed between the choice options, but only container value was consistent with the relative amount of food in the choice options. Finally, in *Equal* trials, container number and container value both differed between the choice options, but the amount of food in each option was the same. We presented the 30 possible combinations of individual containers and homogeneous sets of 2 containers that fit these descriptions.

We presented a total of 86 trials over 4 sessions, with some combinations presented more frequently so that we could have a larger corpus of trials in the critical conditions in which animals had to accommodate both container number and container value to maximize the pellets they could obtain.

Phase 3: One or Two Container(s) vs. Visible Set: In each trial of this phase, monkeys compared 1 or 2 containers of a single type to a set of 1 to 12 visible food items. In trials involving only one container, the container and set of visible items each contained 1 to 5 food items. So, the magnitude of choice options was the same as in Phase 1, with the exception that in some trials, the container and visible set contained the same number of food items, allowing us to assess whether monkeys were biased to respond on the basis of visibility alone. Monkeys experienced all 25 possible comparisons of single containers and sets of visible food items within this range. Monkeys completed a total of 130 trials involving these comparisons across 4 sessions.

In trials involving two containers, we presented monkeys with a more limited selection of comparisons. We selectively included comparisons on the basis of difficulty (magnitude difference), as monkeys' performance on the most difficult trials would be the most informative. Across five sessions, we presented each monkey with 100 trials, of which only 4 consisted of comparisons differing by more than 4 food items, 7 consisted of sets differing by 4 food items, 12 consisted of sets differing by 3 food items, 21 consisted of sets differing by 2 food items, and 36 consisted of sets differing by only 1 food item. The remaining 20

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

trials consisted of comparisons between sets that represented equal food amounts. See Table 2 for examples of each trial type in this phase.

Results

Phase 1: One Container vs. One Container: Liam and Wren exceeded chance performance in their first 40-trial session (binomial sign tests conducted on the number of correct and incorrect responses, both $p = 0.002$), and met the 85% success criterion in 1 and 3 sessions, respectively. Lily and Nala exceeded chance performance in their second and third sessions (both $p < 0.001$), respectively, and they both met the success criterion in their third session. Thus, all 4 capuchins required 120 trials or less to learn the relative amount of food contained in the 5 different containers.

Phase 2: One or Two Container(s) vs. Two Containers: The monkeys performed similarly to one another across the 6 trial types (Figure 1). All 4 monkeys performed nearly perfectly in the 3 trial types in which container value was a relevant cue, *Concordant*, *Discordant – Value*, and *Value* trials (binomial sign tests, all $p < 0.001$). However, few individuals were successful when container number was the only relevant cue. In *Number* trials, only Liam and Wren exceeded chance performance (Liam & Wren: $p = 0.013$; Lily: $p = 0.424$; Nala: $p = 0.791$), and in *Discordant – Number* trials, all 4 monkeys performed either at chance level or significantly below chance level (Liam: $p = 0.021$; Lily: $p < 0.001$; Nala: $p = 0.21$; Wren: $p = 0.077$). In *Equal* trials, all 4 monkeys chose on the basis of container value (Liam & Nala: $p < 0.001$; Lily: $p = 0.011$; Wren: $p = 0.002$).

Phase 3: One or Two Container(s) vs. Visible Set: All four monkeys exceeded chance performance when the visible set was the correct selection, regardless of whether one or two containers made up the alternative choice option (Figure 2a; binomial sign tests, all $p < 0.001$). However, two of the monkeys exceeded chance performance when a single container was the correct selection (Liam: $p = 0.006$; Lily: $p < 0.001$; Nala: $p = 0.366$; Wren: $p = 0.052$). Further, all monkeys performed at or significantly below chance level when two containers were the correct selection (Figure 2b; Liam & Lily: $p = 0.332$; Nala & Wren: $p < 0.001$). When comparing choice options that contained an equal number of food items, most monkeys showed a bias to select the visible set over two containers (Lily: $p = 0.022$. Nala & Wren: $p = 0.002$; Liam: $p = 0.109$), but not over one container (Liam, Lily & Wren: $p = 1.000$; Wren: $p = 0.227$).

Discussion

The capuchins quickly learned the ordinal relations between the 5 individual container types, and they transferred their knowledge of the container values to comparisons of multiple containers. However, few monkeys learned to use container number as a profitable cue, and this prevented the monkeys from effectively summing the amount of food held by multiple containers in a set and using that knowledge to make accurate quantity judgments. The capuchin monkeys' performance in these phases was similar to that of the chimpanzees and macaque previously tested in this paradigm, as they too failed to sum the amount of food held within a container set (Beran et al., 2005).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Those previously tested animals showed a greater level of success when comparing containers to visible food items. They consistently selected the choice option with more food regardless of whether that option consisted of containers or visible food items. This indicated that, not only did those primates understand the relative value of each container, they also understood the cardinal (i.e., absolute) value of each of those containers. The capuchin monkeys tested in the current experiment, however, performed consistently well only when the visible set contained more food. Only half of the monkeys succeeded at this task when the correct selection was a single container, and none did so when the correct selection consisted of two containers. This suggested that only half of these monkeys had some understanding of the cardinality of the containers, and that understanding was limited to simpler comparisons involving only one container.

This deviation in the capuchin monkeys' performance from the previously tested primates seemed to be driven by the formation of a bias to select the visible set of food items. It is possible that this bias stemmed from a more general problem of inhibiting responses to visibly present food items. Such inhibitory issues have been documented in animals in other cognitive paradigms, most notably the reversed-reward contingency task devised by Boysen and her colleagues (e.g., Boysen & Bernston, 1995; for a review of research conducted with this paradigm see Shifferman, 2009). However, these monkeys did not exhibit this bias in previous experiments in which they compared entirely visible food sets to food sets presented item-by-item into an opaque container (Evans et al., 2009). This explanation is also problematic because the monkeys were only biased to select the visible food items when they were presented in comparison to multiple containers (and this pattern was the same even when the visible food option consisted of only a single food pellet). A more likely explanation for this bias is that monkeys became overwhelmed by the introduction of multiple-container choice options to the already difficult test scenario involving comparisons of visible and non-visible choice options, and therefore, defaulted to selecting the visible set of food items.

Experiment 2: Can capuchins and chimpanzees perform ordinal and summative judgments of tokens representing different quantities of discrete food items?

We used a similar method to that employed by Addessi et al. (2007) so that we could compare the performance of chimpanzees to that of capuchins (including the capuchins tested by Addessi et al. (2007) and another sample of capuchins at the Language Research Center). We did vary some of the methodological details, however, to better fit our testing environment. We hypothesized that the chimpanzees and capuchins tested in this experiment would perform comparably to the capuchins previously tested by Addessi et al. (2007) by showing success in the summation trials in this paradigm.

Methods

Participants—We tested the same 4 capuchins immediately following the completion of Experiment 1. We also tested 4 chimpanzees (*Pan troglodytes*) while the monkeys were completing Experiment 1. The chimpanzees were Lana (female, 37 years old), Mercury

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(male, 22 years old), Panzee (female, 23 years old), and Sherman (male, 35 years old). Sherman and Lana previously participated in a nearly identical experiment to Experiment 1 of the present study (Beran et al., 2005), and all four chimpanzees had participated in a variety of other experiments involving judgments between visible and nonvisible sets of food items (e.g., Beran, 2001; Beran & Beran, 2004). The chimpanzees also participated in one previous study involving tokens (Brosnan & Beran, 2009).

It is important to note that three of the four chimpanzees (all but Mercury) were raised in an environment in which they could use lexigram symbols and respond to the symbols used by others (see Rumbaugh and Washburn, 2003). Lana's interaction with humans was via an electronic keyboard on which she would have to type lexigram sentences to ask for various foods, request certain activities, or to answer questions addressed to her by her human caretakers (Rumbaugh, 1977). Sherman was raised with another chimpanzee and learned to use lexigrams and respond to lexigram use by other chimpanzees and by humans (Savage-Rumbaugh, 1986). Panzee was co-reared with a bonobo in a highly enriched environment that led to her use lexigrams and to respond to spoken English (Brakke & Savage-Rumbaugh, 1995, 1996). Thus, these chimpanzees have demonstrated high levels of symbol use and comprehension in both linguistic and other situations, and such skills may play a role in the performance of these chimpanzees in the experiments described in this paper.

Design and Procedure—We used a similar procedure to Experiment 1 to present token comparisons to the capuchins. We again tested individual monkeys by requiring them to reach through 1 of 2 holes in their test enclosures to touch (but not take) 1 of 2 choice options presented on top of a moveable utility cart. This was different from the Addessi et al. (2007) method, but we used this procedure because it was highly familiar to the monkeys and we did not expect there to be any advantage to having the monkeys handle the tokens any more than this (e.g., by handing the tokens to the experimenter, as was done in the original study). To test the chimpanzees, we used a previously established procedure in which individuals reached their fingers through the wire mesh of their enclosure to touch 1 of 2 choice options presented on a sliding shelf (as in Beran et al., 2005). In half of the test sessions, the chimpanzees did not handle the selected token(s) (similar to the capuchin procedure), and in the other half of the test sessions they received the selected token(s) from one experimenter and transferred them to a second experimenter in a nearby location (see below for more details). As tokens, we used rod-shaped pieces of hard white plastic (polypropylene) and aluminum to represent 1 and 3 food items (hereafter referred to as *one-item* and *three-item* tokens), respectively.[1] The token dimensions were 1.25 cm × 1.25 cm × 2.5 cm for the capuchins and 1.25 cm × 1.25 cm × 10 cm for the chimpanzees. We presented the tokens, side-by-side, in 2 large bowls similarly to how we presented containers in Experiment 1. We rewarded the monkeys' token selections with the appropriate number of 45-mg banana-flavored pellets. The monkeys were accustomed to working for pellet rewards and had proven highly motivated to obtain such rewards in previous studies and in Experiment 1 of the present study. We rewarded the chimpanzees' token selections with a

[1]One chimpanzee (Sherman) initially exhibited an aversion to the aluminum rods used as three-item tokens during the training phase. So, for this individual only, we introduced similar-sized pieces of gray PVC pipe as three-item tokens. This eliminated his avoidance of the three-item token, and thus, we used this material with this subject for all other trials in this experiment.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

variety of nuts, cereals, candies, and fruit pieces, though we only varied food type between sessions (i.e., the same type of food was used in association with both token types for entire sessions). We rewarded the chimpanzees with a greater variety of items to keep them interested in the task, as they were accustomed to working for higher quality reward items than manufactured food pellets.

Phase 1: Single One-Item Token vs. Single Three-Item Token: Because we required only the chimpanzees to actually handle and exchange the tokens, and because the chimpanzees were already familiar with this procedure (Brosnan and Beran, 2009), we did not conduct separate training sessions of entirely forced-choice trials, as was done by Addessi et al. (2007). Instead, we presented the chimpanzees and monkeys with alternating blocks of forced-choice trials and free-choice trials within the same session. Forced-choice trials presented only one choice option, either a single one-item token or a single three-item token. Free-choice trials presented two choice options, both a single one-item token and a single three-item token. In each session, we presented a total of 30 trials consisting of alternating blocks of 4 forced-choice trials and 6 free-choice trials. We conducted one session with each individual, no more than once per day, until they reached a success criterion of 85% correct free-choice trials within a single session. This criterion was slightly different than what was used in the original study (90% correct exchanges in two consecutive training sessions), but because we included a small block of forced-choice trials at the beginning of each test session, we believed that this training criterion would not influence later performance.

The training procedure differed slightly between species, in that the chimpanzees were actually handed the tokens, after which they would transport the tokens to an adjacent cage and hand them to a second experiment to receive the appropriate amount of food items. We believed that this would make the task more interesting for the chimpanzees, and thus, would make them more likely to participate on a continual basis. We did not do this with the capuchins because we had no area through which they could move in a similar way. Also, capuchins were not handed tokens because our general training method with these animals had always involved discrete responses and animals had been trained not to take things from the test apparatus.

Phase 2: Single Three-Item Token vs. Multiple One-Item Tokens: We presented both species with the exact same test trial types used in Experiment 1 of the original study by Addessi et al. (2007) with capuchins. In each test trial, a single three-item token was compared to 1 to 5 one-item tokens, which made 5 different types of test trials. We also included a small block of forced-choice trials at the beginning of each test session to remind the animals of the value of individual choice options. There was only one available choice option in these forced trials, and it consisted of 1 one-item token, 1 three-item token, 2 one-item tokens, 3 one-item tokens, or 4 one-item tokens. A single test session consisted of 5 forced-choice trials (1 trial of each type above) followed by a block of 10 test comparisons (2 trials of each type). This number of trials was small in comparison to the average test session in which these monkeys and chimpanzees typically participate, so we expected them to be motivated to maximize food intake throughout the session. Each animal completed exactly 20 trials of each free-choice trial type (i.e., 100 total free-choice trials) using this procedure. Each chimpanzee completed an additional 20 trials of each type using a procedure identical to how the monkeys responded (i.e., the 'touch but don't take' method), so we could assess the influence of the general test procedure on the chimpanzees' performance.[2] We hereafter refer to these two blocks of sessions completed by the chimpanzees as Phase 2a (in which they handed tokens to an experimenter) and Phase 2b (in which they only touched the token sets through their cage wire).

Phase 3: One or Two Three-Item Tokens vs. Multiple One-Item Tokens: We tested only the chimpanzees in this phase, on the basis of a species difference in Phase 2 performance (see Results). We presented the chimpanzees with the exact same test trial types used in Experiment 2 of the Addessi et al. (2007) study, as well as one new trial type that we believed would be informative (Table 3). The new trial type involved a comparison in which choosing 2 three-item tokens would be an incorrect response. We added this trial type to discourage the formation of the rule "always choose the two three-item tokens." As in Phase 2, we also included a small block of forced-choice trials at the beginning of each test session to remind the animals of the value of individual choice options. These forced trials included only one available choice option and it consisted of 1 one-item token, 1 three-item token, 2 one-item tokens, or 2 three-item tokens. A single session consisted of 4 forced-choice trials (1 trial of each type) followed by a block of 14 test comparisons (2 trials of each type). As stated above, this number of trials was small in comparison to the chimpanzees' average test session, so we expected them to be motivated to maximize food intake throughout the session. Each chimpanzee completed exactly 20 trials of each free-choice trial type for a total of 140 free-choice trials. Results

Phase 1: Single One-Item Token vs. Single Three-Item Token: The monkeys Wren and Liam were above chance performance beginning in their 2nd and 3rd training sessions (binomial sign tests, $p = 0.031$, $p < 0.001$), respectively, and they each required 3 sessions to reach the 85% success criterion regarding free-choice trials. The monkeys Lily and Nala were above chance level of performance in their 2nd and 4th sessions ($p = 0.008$, $p = 0.031$), respectively, and required 7 and 6 training sessions to meet the success criterion. Therefore, each monkey required no more than 126 free-choice trials to complete the training phase.

The chimpanzees Lana and Panzee performed above chance level (binomial sign tests, both $p < 0.001$) and met the 85% success criterion in their first training session. The chimpanzees Mercury and Sherman were above chance performance level in their 2nd and 5th sessions ($p = 0.031$; $p < 0.001$), respectively, and required 3 and 5 sessions to meet the criterion.[3] Thus, each chimpanzee required no more than 90 free-choice trials to complete the training phase.

Phase 2: Single Three-Item Token vs. Multiple One-Item tokens: The capuchins showed a nearly complete bias to select the three-item token in all trial types, whether or not it was the correct choice. They were, therefore, significantly above chance performance in all comparisons involving 1 or 2 one-item tokens (binomial sign tests, all $p < 0.001$), and

[2]There was one exception to this: Panzee became ill near the end of Phase 2b, and thus, was not able to complete as many trials as the other chimpanzees. All analyses regarding Panzee's performance in this half of the phase were based on 15 trials in each category.
[3]More accurately, Sherman required 4 sessions with the original token materials (plastic and aluminum) and one session with his new token material (gray PVC in place of aluminum) to reach the training success criterion.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

significantly below chance performance in all comparisons involving 4 or 5 one-item tokens (all $p < 0.001$).

The chimpanzees were more variable in performance between individuals (Figure 3). All 4 chimpanzees performed significantly above chance level in comparisons involving 1 one-item token and 1 three-item token, both in Phase 2a when they handled and transported the tokens (binomial sign tests; Lana: $p = 0.041$; Mercury: $p = 0.012$; Panzee and Sherman: both $p < 0.001$; Figure 3a – first bar cluster), and in Phase 2b when they merely pointed to the tokens (Lana: $p < 0.001$; Mercury: $p = 0.003$; Panzee and Sherman: both $p < 0.001$; Figure 3b – first bar cluster). However, Lana and Mercury exhibited a bias to select the set of one-item tokens in the remainder of trial types in Phases 2a and 2b, making them significantly below chance performance in all other comparisons involving 2 one-item tokens (Figure 3a & b – second bar cluster), and significantly above chance performance in all comparisons involving 4 or 5 one-item tokens (all $p < 0.001$; Figure 3a & b – last two bar clusters).

Panzee and Sherman were more successful, as they were significantly above chance performance in trials involving 2 one-item tokens in either Phase 2a (Panzee: $p = 0.003$; Sherman: $p = 0.011$; Figure 3a – second bar cluster) or Phase 2b (Panzee: $p = 1.000$; Sherman: $p = 0.041$; Figure 3b – second bar cluster). They also performed significantly above chance levels on comparisons involving 4 one-item tokens in either Phase 2a (Panzee: $p = 0.263$; Sherman: 2a – $p = 0.215$; Figure 3a – fourth bar cluster) or Phase 2b (Panzee: $p = 0.007$; Sherman: $p = 0.041$; Figure 3b – fourth bar cluster), as well as on comparisons involving 5 one-item tokens in either Phase 2a (Panzee: $p = 0.607$; Sherman: $p = 0.263$; Figure 3a – fifth bar cluster) or Phase 2b (Panzee: $p = 0.007$; Sherman: $p = 0.003$; Figure 3b – fifth bar cluster).

Phase 3: One or Two Three-Item Tokens vs. Multiple One-Item Tokens: Lana and Mercury continued to show a bias to select the set of multiple one-item tokens, regardless of the number of tokens in that set, or the number of three-item tokens in the comparison set. This made them both significantly above chance performance when comparing 1 three-item token to 4 one-item tokens (binomial sign tests; Lana: $p < 0.001$; Mercury: $p = 0.003$; Figure 4a – second bar cluster) or 5 one-item tokens (both $p < 0.001$; Figure 4a – third bar cluster), and when comparing 2 three-item tokens to 8 one-item tokens (both $p < 0.001$; Figure 4b – fourth bar cluster). This also made them significantly below chance performance when comparing 2 three-item tokens to 4 one-item tokens (Lana: $p < 0.001$; Mercury: $p = 0.019$; Figure 4b – first bar cluster) or 5 one-item tokens (Lana: $p < 0.001$; Mercury: $p = 0.003$; Figure 4b – second bar cluster).

Panzee and Sherman did not exhibit a bias to select one token type over the other, but they also did not often select the token set representing the greatest quantity of food items. Unlike in Phase 2, these chimpanzees did not consistently select 4 one-item tokens (Panzee: $p = 0.115$; Sherman: $p = 0.263$; Figure 4a – second bar cluster) or 5 one-item tokens (Panzee: $p = 0.003$; Sherman: $p = 0.503$; Figure 4a – third bar cluster) over 1 three-item token. With one exception, they also performed at chance levels when comparing 2 three-item tokens to either 4 one-item tokens (Panzee: $p = 0.263$; Sherman: $p < 0.001$; Figure 4b – first bar

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript

cluster), 5 one-item tokens (both $p = 0.503$; Figure 4b – second bar cluster), or 8 one-item tokens (both $p = 0.503$; Figure 4b – fourth bar cluster).

Discussion

Both the capuchins and the chimpanzees learned the relative amount of food associated with choosing each token type in this experiment with the same speed that they acquired the relative container values in Experiment 1 (capuchins) and the previously conducted container study (chimpanzees; Beran et al., 2005). Also, all members of both species transferred their ability to select a single three-item token over a single one-item token to Phase 2. However, in the remaining trial types of Phase 2, only half of the chimpanzees (Panzee and Sherman) and none of the capuchin monkeys performed at the same level as the capuchins reported in Experiment 1 of Addessi et al. (2007), by selecting the single three-item token or set of one-item tokens representing the greater amount of food. In the final phase, in which trials involving 2 three-item tokens were introduced, no animals were successful in more than one trial type, and thus showed no consistent capacity to sum and compare the token sets presented in both choice options. Therefore, only two chimpanzees exhibited any level of token summation in this experiment, and it was limited to simpler trials involving a single three-item token.

One potential explanation for the primates' Phase 2 and Phase 3 performance is a lack of motivation to maximize food acquisition, given our inclusion of rewarded forced trials at the beginning of each session and the small differences in food amounts that we presented in choice trials. However, all of these primates have maintained above-chance performance levels in other quantity choice studies yielding similar overall food amounts and involving similar numerical distances between choice options (e.g., Beran, 2001; Beran et al., 2008a; Evans et al., 2009). The same explanation would have applied in those experiments as well, and yet those primates consistently performed well throughout sessions for differences in reward amounts as small as in the present experiment. Therefore, this was most likely not the reason for their decline in performance.

It is important to note that there were differences in the method used here and that used by Addessi et al. (2007), and those differences may have influenced performance levels. For example, the monkeys in the original study were presented with eight times as many trials in each of the critical test phases as the monkeys and apes in the present study, though some of those monkeys had reached statistically significant performance on several of the trial types in the same number of trials presented in the current experiment. Also, the animals tested in the present study had different levels of contact with the tokens than did the animals originally tested by Addessi et al. (2007). It is possible that the elevated amount of direct contact the previously tested capuchin monkeys had with the tokens resulted in their greater learning and performance in test trials. Similarly, the small differences in the chimpanzees' performance between Phases 2a and 2b may have been the result of the different levels of handling requirements in those phases. Panzee and Sherman performed better in some trial types of Phase 2b, in which they simply touched one of two token sets to receive rewards, in comparison to the same trial types in Phase 2a, in which they had to walk to an adjacent enclosure and trade the tokens for rewards. This seemed to indicate that the requirement of

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

moving from one enclosure to another and the action of handing out tokens in Phase 2a was negatively influencing the chimpanzees' performance. However, because we did not control the order of these phases, it is possible that it was the added experience of the second phase, and not the difference in methodology, that created the difference in performance level.

General Discussion

Overall, this study provided only limited evidence of the summation of symbolic stimuli by either chimpanzees or capuchins except when they could rely on certain cues (such as the highest value token or symbol being a part of the overall largest set). This suggests that flexible combination of symbolic stimuli in quantity judgment tasks may be within the abilities of chimpanzees and capuchins but does not characterize the majority of individuals. However, these results also indicate the need to carefully examine specific methodological details that may promote or hinder such possible summation.

Previously, chimpanzees showed no ability to compare two sets of symbolic stimuli for quantity in which the overall larger quantity consisted of lower value stimuli of a greater number (Beran et al., 2005). Experiment 1 of this study extended that limitation to capuchins, and the similarities in performance between capuchins and chimpanzees were rather striking. Both species learned ordinal relations between symbols for quantity at approximately the same rate. Both species were capable to some degree of comparing symbols of quantity to actual food quantities in a judgment task, although capuchins were more limited in this regard compared to chimpanzees and a macaque. However, success at any level indicated that these symbols did represent real quantities for the animals. Being able to compare these symbols to visible quantities also indicated that these symbols did represent absolute quantities rather than only relative quantities (see Beran et al., 2005, for more discussion of this point).

A different method that involved only two token types did produce successful performance in four out of 10 capuchins in a previous study (Addessi et al., 2007). However, the present study failed to produce the same outcome with another sample of capuchins. None of the latter capuchins performed in a way that indicated they combined token values in flexible ways to obtain more as opposed to less food. Rather, capuchins showed biases for specific tokens of a given value, indicating that they did not accommodate both the value of the tokens in a set and the number of those tokens in that set. Because the capuchins tested by Addessi et al. (2007) were already successful in some test comparisons after the same number of trials presented in the present study, it is likely that individual differences accounted for the contrasting results. This hypothesis is supported by the fact that two of four chimpanzees did show some ability to combine token values in Experiment 2, although this performance was limited. In some cases, these two chimpanzees selected the overall larger amount of food, but such performance was not sustained throughout the experiment. However, it is possible that providing these chimpanzees with the same degree of experience received by the capuchins tested by Addessi et al. (2007) might have improved their performance. Other aspects of the experiments in the present study also differed from those of previous studies (especially Addessi et al., 2007), and so these are important to remember. Although we believed that the differences were not substantial, and that the new methods

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

were at least conceptually related to the previous studies, this may not be true. It may be that the differences in methodology were relevant to the different outcomes by preventing the tasks from requiring the same kinds of representational abilities. Our tests, and those of Addessi et al. (2007), may be valid assessments of the symbol summation abilities of primates, but they may require different response mechanisms, some of which are available to chimpanzees and capuchins and some of which are not.

There is precedent to expect that such summation performance is possible. Boysen and Berntson (1989) reported that the chimpanzee Sheba could move to separate locations where different Arabic numerals were presented and then return to a start location and provide the numeral that indicated the sum of the previous two numerals. Sheba had to represent the summed total of the items she had viewed to choose the correct numeral. There is also some evidence that animals can compare two sets of stimuli that have accrued an association with specific quantities of food and choose the overall larger set (Olthof et al., 1997; Olthof & Roberts, 2000). However, those tasks are different than the task presented in Experiment 1 and by Beran et al. (2005) in which each choice set was composed of two different kinds of information: a container/egg of a given value, and the number of those container/eggs in the set. This required a different form of summation, and one that Beran et al. (2005) argued was in some way analogous to a simple multiplication problem with its attendant problem that successful understanding of its product requires that one not focus on only a single factor (e.g., knowing that 9*2 is not more than 6*5 even though 9 is the biggest individual number). When given this kind of trial, chimpanzees failed completely. In contrast, the task developed by Addessi et al. (2007), and used here in Experiment 2, does not rely as heavily on the relative value component as there are only two token types (instead of the five used in Experiment 1). This may account for some of the behavioral differences with the greater likelihood of success (although, again, Experiment 2 provided only limited evidence of such success for these chimpanzees and capuchins).

Another possible interpretation of these studies is that nonhuman primates' ability to track and judge relative quantity is hindered when comparison quantities are hidden or represented by a symbol. Some studies involving spontaneous judgments between single non-visible quantities have reported only limited success with such comparisons, with performance waning for quantities beyond 3 or 4. Such findings have been shown for experimentally naïve monkeys (e.g., Hauser & Carey, 2003; Hauser, Carey & Hauser, 2000), and replicated with human infants (e.g., Feigenson, Carey & Hauser, 2002). However, there is a larger body of evidence with more experienced nonhuman primates indicating that they can make relative quantity judgments between single non-visible quantities of greater magnitude, such as sets of 1 to 9 items (e.g., Beran, 2001, 2007; Beran & Beran, 2004; Beran et al., 2008a; Beran & Beran, 2004; Evans et al., 2009; Hanus & Call, 2007). So, it is more likely that the performance pattern seen in the present study and the studies on which it is based is due to an inability to sum quantities represented by multiple tokens, not simply an inability to represent and compare single out-of-view quantities represented by such tokens.

What seems clear is that container value is more heavily weighted in the overall representation of quantity in a set composed of multiple containers of the same value. Number of containers appears to be less salient in the representation of total quantity. This

Author Manuscript

effect matches other studies where different kinds of cues to a correct response were offered to primates, and some cues were more profitable than others. For example, Call (2003) reported that apes were better at tracking the movement of hidden items when the container with those items was indicated by showing the items themselves rather than showing a marker sitting on top of the container. Here, the difference was not simply the result of markers being less useful as a cue than the food itself, because apes performed equally well with both cues when they only had to remember the location without having to also track (an easier task in general; see also Okamoto-Barth & Call, 2008). Beran, Evans, and Harris (2008) reported that chimpanzees made suboptimal responses when presented with two sets of food items of the same type by choosing sets with the largest individual piece of food but the smaller total amount. Here, the bias was to use item size, and not total food amount, in making judgments. Boysen, Berntson, and Mukobi (2001) reported that chimpanzees used the size of individual food items more heavily in choosing between sets of candies that differed in size and quantity, and this led to choices of overall smaller amounts of food in some cases. Beran, Evans, and Ratliff (2009) also reported that chimpanzees undervalued the total amount of food in sets when items differed in size and did not appear to be whole. Here, the bias was to select sets of food that did not appear to be fractionated, even though those sets actually contained less food overall. All of these results indicate that decision-making by primates in the face of multiple cues sometimes differs from optimality because of biases or inability to use the most relevant information presented within trials.

In summary, this study offers three main points regarding symbolic representation of quantity by nonhuman primates. First, several species (capuchins, macaques, and chimpanzees) can readily and quickly learn to associate arbitrary stimuli with specific quantities and then compare those stimuli in the same way that they compare visible quantities. Second, these species treat such symbols as representations with values equivalent to real food items, as they can successfully compare individual and (in some cases) pairs of symbols to visible sets of food items. Third, there appears to be some limitation in the capacity of nonhuman primates to appropriately represent and accommodate both the individual value of tokens in a given set and the number of tokens in that set, in a way that allows for flexible and accurate representations of quantity for the purpose of comparing and choosing between such sets. This may be a limit to the calculation and estimation capacities of nonhuman animals, although there are individual differences and further methodological variations may yet produce more compelling positive evidence.

Acknowledgments

This research project was supported by grant HD-38051 from the National Institute of Child Health and Human Development and grants SES – 0729244 and BCS - 0729244 from the National Science Foundation.

References

Addessi E, Crescimbene L, Visalberghi E. Do capuchin monkeys (*Cebus apella*) use tokens as symbols? Proceedings of the Royal Society of London. 2007; 274:2579–2585.

Addessi E, Crescimbene L, Visalberghi E. Food and quantity token discrimination in capuchin monkeys (*Cebus apella*). Animal Cognition. 2008a; 11:275–282. [PubMed: 17901990]

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

Addessi E, Mancini A, Crescimbene L, Padoa-Schioppa C, Visalberghi E. Preference transitivity and symbolic representation in capuchin monkeys (*Cebus apella*). PLoS ONE. 2008b; 3:e2414. [PubMed: 18545670]

Beran MJ. Summation and numerousness judgments of sequentially presented sets of items by chimpanzees (*Pan troglodytes*). Journal of Comparative Psychology. 2001; 115:181–191. [PubMed: 11459165]

Beran MJ. Rhesus monkeys (*Macaca mulatta*) enumerate large and small sequentially presented sets of items using analog numerical representations. Journal of Experimental Psychology: Animal Behavior Processes. 2007; 33:42–54. [PubMed: 17227194]

Beran MJ. Capuchin monkeys (*Cebus apella*) succeed in a test of quantity conservation. Animal Cognition. 2008; 11:109–116. [PubMed: 17549530]

Beran MJ, Beran MM. Chimpanzees remember the results of one-by-one addition of food items to sets over extended time periods. Psychological Science. 2004; 15:94–99. [PubMed: 14738515]

Beran MJ, Beran MM, Harris EH, Washburn DA. Ordinal judgments and summation of nonvisible sets of food items by two chimpanzees and a rhesus macaque. Journal of Experimental Psychology: Animal Behavior Processes. 2005; 31:351–362. [PubMed: 16045389]

Beran MJ, Evans TA, Ratliff CL. Perception of food amounts by chimpanzees (*Pan troglodytes*): The role of magnitude, contiguity, and wholeness. Journal of Experimental Psychology: Animal Behavior Processes. 2009; 35:516–524. [PubMed: 19839704]

Beran MJ, Evans TA, Harris EH. Perception of food amount by chimpanzees based on the number, size, contour length, and visibility of items. Animal Behaviour. 2008; 75:1793–1802. [PubMed: 19412322]

Beran MJ, Evans TA, Leighty KA, Harris EH, Rice D. Summation and quantity judgments of sequentially presented sets by capuchin monkeys (*Cebus apella*). American Journal of Primatology. 2008a; 70:191–194. [PubMed: 17879377]

Beran MJ, Harris EH, Evans TA, Klein ED, Chan B, Flemming TM, Washburn DA. Ordinal judgments of symbolic stimuli by capuchin monkeys (*Cebus apella*) and Rhesus monkeys (*Macaca mulatta*): The effects of differential and nondifferential reward. Journal of Comparative Psychology. 2008b; 122:52–61. [PubMed: 18298281]

Beran MJ, Rumbaugh DM. “Constructive” enumeration by chimpanzees (*Pan troglodytes*) on a computerized task. Animal Cognition. 2001; 4:81–89.

Boysen ST, Berntson GG. Numerical competence in a chimpanzee (*Pan troglodytes*). Journal of Comparative Psychology. 1989; 103:23–31. [PubMed: 2924529]

Boysen ST, Berntson GG. Responses to quantity—perceptual versus cognitive mechanisms in chimpanzees (*Pan troglodytes*). Journal of Experimental Psychology: Animal Behavior Processes. 1995; 21:82–86. [PubMed: 7844508]

Boysen ST, Berntson GG, Hannan MB, Cacioppo JT. Quantity-based interference and symbolic representations in chimpanzees (*Pan troglodytes*). Journal of Experimental Psychology: Animal Behavior Processes. 1996; 22:76–86. [PubMed: 8568498]

Boysen ST, Berntson GG, Mukobi KL. Size matters: Impact of item size and quantity on array choice by chimpanzees (*Pan troglodytes*). Journal of Comparative Psychology. 2001; 115:106–110. [PubMed: 11334213]

Boysen ST, Mukobi KL, Berntson GG. Overcoming response bias using symbolic representations of number by chimpanzees (*Pan troglodytes*). Animal Learning & Behavior. 1999; 27:229–235.

Boysen ST, Hallberg KI. Primate numerical competence: Contributions toward understanding nonhuman cognition. Cognitive Science. 2000; 24:423–443.

Brakke KE, Savage-Rumbaugh ES. The development of language skills in bonobo and chimpanzee: I. Comprehension. Language and Communication. 1995; 15:121–148.

Brakke KE, Savage-Rumbaugh ES. The development of language skills in Pan: II. Production. Language and Communication. 1996; 16:361–380.

Brannon EM, Terrace HS. Representation of the numerosities 1–9 by rhesus macaques (*Macaca mulatta*). Journal of Experimental Psychology: Animal Behavior Processes. 2000; 26:31–49. [PubMed: 10650542]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Brosnan SF, Beran MJ. Trading behavior between conspecifics in chimpanzees, *Pan troglodytes*. Journal of Comparative Psychology. 2009; 123:181–194. [PubMed: 19450025]

Brosnan SF, De Waal FBM. Monkeys reject unequal pay. Nature. 2003; 425:297–299. [PubMed: 13679918]

Brosnan SF, De Waal FBM. A concept of value during experimental exchange in brown capuchin monkeys, *Cebus apella*. Folia Primatologica. 2004; 75:317–330.

Brosnan SF, Schiff HC, de Waal FBM. Tolerance for inequity may increase with social closeness in chimpanzees. Proceedings of the Royal Society of London, Series B. 2005; 1560:253–258.

Brosnan SF, Jones OD, Lambeth SP, Mareno MC, Richardson AS, Schapiro SJ. Endowment effects in chimpanzees. Current Biology. 2007; 17:1704–1707. [PubMed: 17884499]

Call J. Spatial rotations and transpositions in orangutans (*Pongo pygmaeus*) and chimpanzees (*Pan troglodytes*). Primates. 2003; 44:347–357. [PubMed: 12937996]

Chen MK, Lakshminaryanan V, Santos LR. The evolution of our preferences: Evidence from capuchin monkey trading behavior. Journal of Political Economy. 2006; 114:517–537.

Deacon, T. The Symbolic Species: The Co-Evolution of Language and the Brain. New York & London: W. W. Norton & Co; 1997.

Evans TA, Beran MJ, Harris EH, Rice DF. Quantity judgments of sequentially presented food items by capuchin monkeys (*Cebus apella*). Animal Cognition. 2009; 12:97–105. [PubMed: 18670794]

Feigenson L, Carey S, Hauser MD. The representation underlying infants' choice of more: Object files versus analog magnitudes. Psychological Science. 2002; 13:150–156. [PubMed: 11933999]

Gardner RA, Gardner BT. Teaching sign language to a chimpanzee. Science. 1969; 165:664–672. [PubMed: 5793972]

Hauser MD, Carey S, Hauser LB. Spontaneous number representation in semi-free-ranging rhesus monkeys. Proceedings of the Royal Society of London, Series B. 2000; 267:829–833. [PubMed: 10819154]

Herman LN, Richards DG, Wolz JP. Comprehension of sentences by bottlenose dolphins. Cognition. 1984; 16:129–219. [PubMed: 6540652]

Hanus D, Call J. Discrete quantity judgments in the great apes (*Pan paniscus, Pan troglodytes, Gorilla gorilla, Pongo pygmaeus*): The effect of presenting whole sets versus item-by-item. Journal of Comparative Psychology. 2007; 121:241–249. [PubMed: 17696650]

Judge PJ, Evans TA, Vyas D. Ordinal representation of numeric quantities by brown capuchin monkeys (*Cebus apella*). Journal of Experimental Psychology: Animal Behavior Processes. 2005; 31:79–94. [PubMed: 15656729]

Lakshminaryanan V, Chen MK, Santos LR. Endowment effect in capuchin monkeys. Philosophical Transactions of the Royal Society of London, Series B. 2008; 363:3837–3844. [PubMed: 18840573]

Matsuzawa T. Use of numbers by a chimpanzee. Nature. 1985; 315:57–59. [PubMed: 3990808]

Mitchell RW, Yao P, Sherman PT, O'Regan M. Discriminative responding of a dolphin (*Tursiops truncatus*) to differentially rewarded stimuli. Journal of Comparative Psychology. 1985; 99:218–225.

Okamoto-Barth S, Call J. Tracking and inferring spatial rotation by children and great apes. Developmental Psychology. 2008; 44:1396–1408. [PubMed: 18793071]

Olthof A, Iden CM, Roberts WA. Judgments of ordinality and summation of number symbols by squirrel monkeys (*Saimiri sciureus*). Journal of Experimental Psychology: Animal Behavior Processes. 1997; 23:325–339. [PubMed: 9206027]

Olthof A, Roberts WA. Summation of symbols by pigeons (*Columba livia*): The importance of number and mass of reward items. Journal of Comparative Psychology. 2000; 114:158–166. [PubMed: 10890587]

Pepperberg IM. Evidence for conceptual quantitative abilities in the African Grey Parrot: Labeling of cardinal sets. Ethology. 1987; 75:37–61.

Pepperberg IM. Numerical competence in an African Grey parrot (*Psittacus erithacus*). Journal of Comparative Psychology. 1994; 108:36–44.

Pepperberg, IM. The Alex studies: Cognitive and communicative abilities of Grey parrots. Cambridge, MA: Harvard University Press; 1999.

Pepperberg IM. Ordinality and inferential abilities of a grey parrot (*Psittacus erithacus*). Journal of Comparative Psychology. 2006; 120:205–216. [PubMed: 16893258]

Premack, D.; Premack, AJ. The Mind of an Ape. New York: Norton; 1983.

Rumbaugh, DM. Language Learning by a Chimpanzee: The Lana Project. New York: Academic Press; 1977.

Rumbaugh DM, Hopkins WD, Washburn DA, Savage-Rumbaugh ES. Lana chimpanzee learns to count by "NUMATH": A summary of a videotaped experimental report. Psychological Record. 1989; 39:459–470. [PubMed: 11540081]

Rumbaugh, DM.; Washburn, DA. Intelligence of Apes and Other Rational Beings. New Haven: Yale University Press; 2003.

Savage-Rumbaugh, ES. Ape Language: From Conditioned Response to Symbol. New York: Columbia University Press; 1986.

Schusterman RJ, Gisiner R. Artificial language comprehension in dolphins and sea lions: The essential cognitive skills. Psychological Record. 1988; 38:311–348.

Schusterman RJ, Krieger K. California sea lions are capable of semantic comprehension. Psychological Record. 1984; 34:3–23.

Shifferman EM. Its own reward: lessons to be drawn from the reversed-reward contingency paradigm. Animal Cognition. 2009; 12:547–558. [PubMed: 19205761]

Tomonaga M, Matsuzawa T. Sequential responding to arabic numerals with wild cards by the chimpanzee (*Pan troglodytes*). Animal Cognition. 2000; 3:1–11.

Washburn DA, Rumbaugh DM. Ordinal judgments of numerical symbols by macaques (*Macaca mulatta*). Psychological Science. 1991; 2:190–193. [PubMed: 11537106]

Westergaard GC, Evans TA, Howell S. Token mediated tool exchange between tufted capuchin monkeys (*Cebus apella*). Animal Cognition. 2007; 10:407–414. [PubMed: 17345056]

Westergaard GC, Liv C, Chavanne TJ, Suomi SJ. Token-mediated tool-use by a tufted capuchin monkey (*Cebus apella*). Animal Cognition. 1998; 1:101–106. [PubMed: 24399274]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

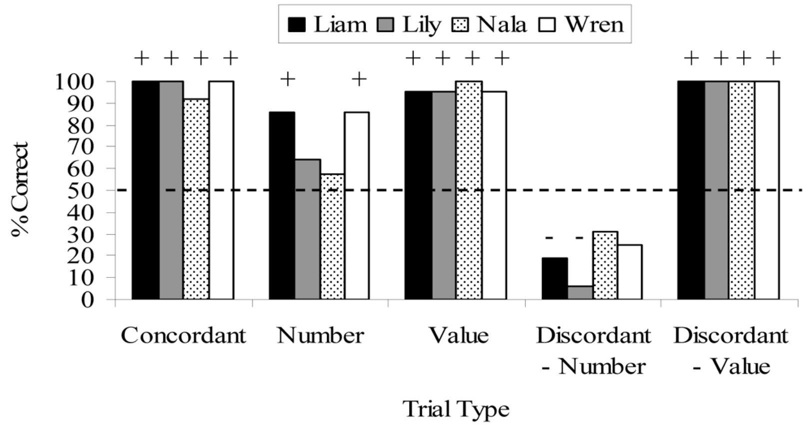
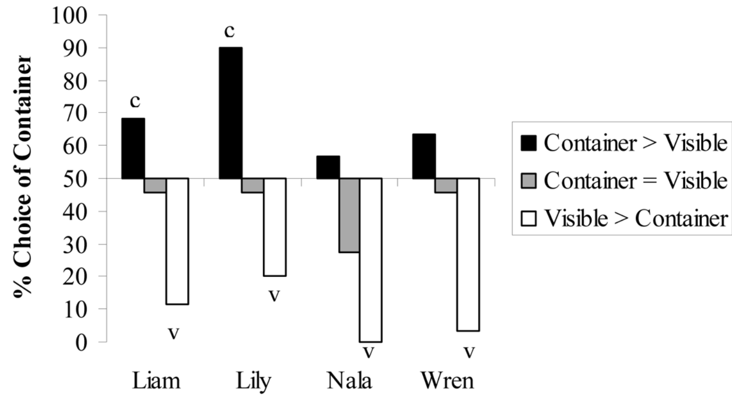


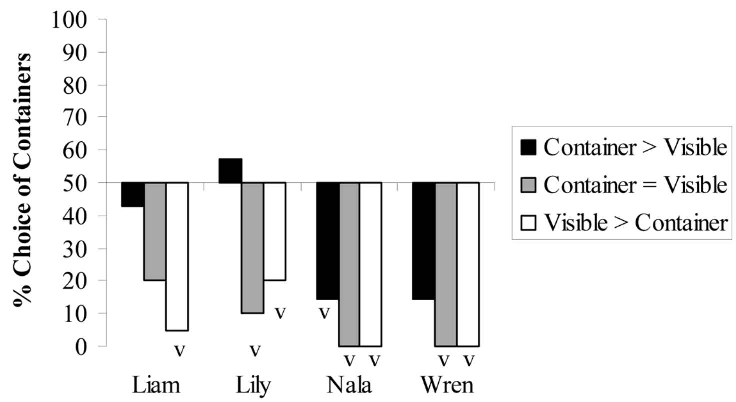
Figure 1.

The monkeys’ performance in Phase 2 of Experiment 1 as a function of trial type. The *Equal* trial type is excluded because there was no correct response in these trials. The horizontal line indicates chance performance. Plus and minus signs represent performance significantly above and below chance level, respectively (two-tailed binomial test, $p < 0.05$).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript



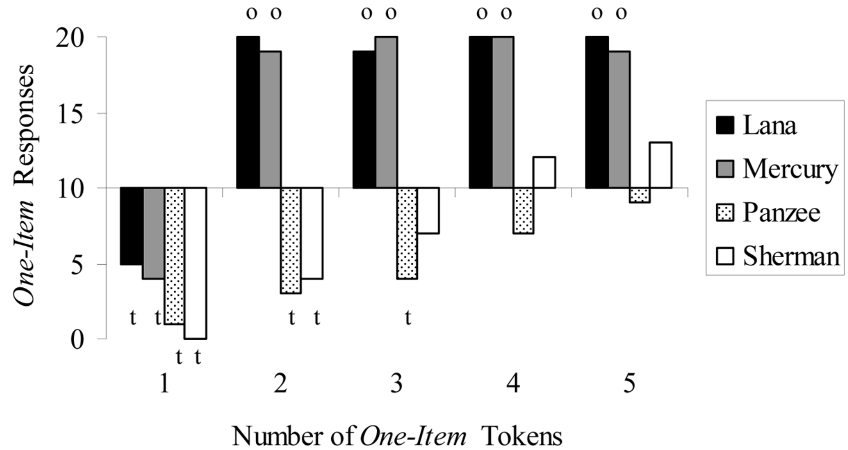
a. One Container vs. Visible Set



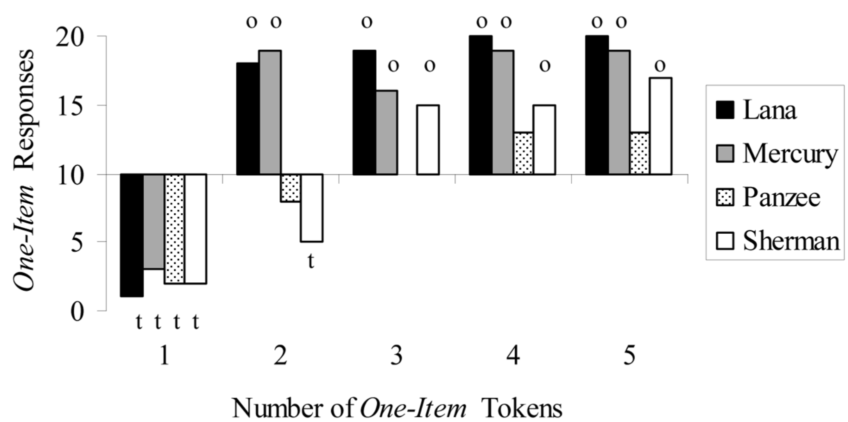
b. Two Containers vs. Visible Set

Figure 2.

The monkeys' performance in Experiment 1, Phase 3 in which visible food items were compared to one container (a) or two food containers (b). The x-axis represents chance performance, so bars above and below the x-axis represent selection biases for the container(s) and visible food items, respectively. Lower case c's and v's denote statistically significant selection biases for the container option and the visible food option, respectively (two-tailed binomial test, $p < 0.05$).



a. Chimpanzees handed the token(s) to an experimenter to receive their reward(s)



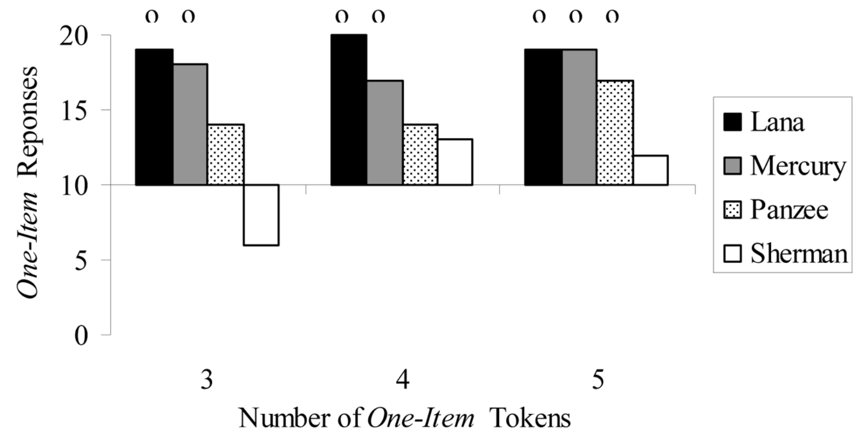
b. Chimpanzees touched the token(s) to receive their reward(s)

Figure 3.

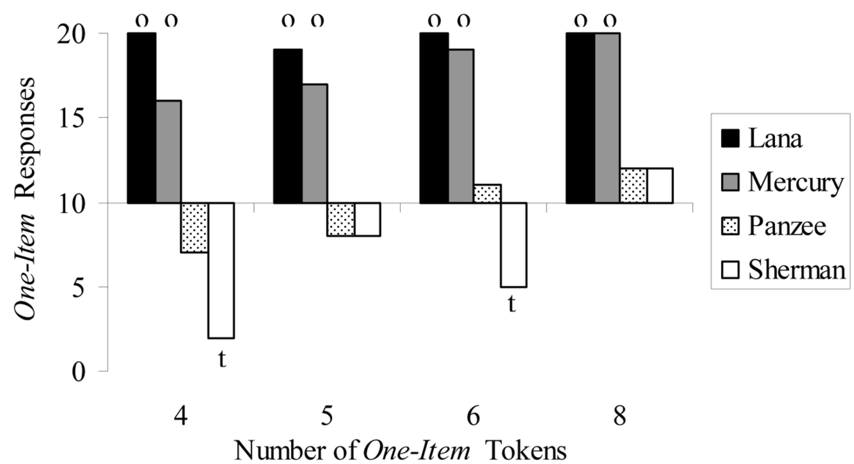
Chimpanzees' performance in Phase 2 of Experiment 1 as a function of the number of *one-item* tokens being compared to a single *three-item* token. Data are split into separate charts on the basis of general test procedure. The x-axis represents the level of chance performance, so bars above and below the x-axis represent selection biases for the one-item token(s) and three-item token, respectively. Lower case o's and t's denote statistically significant biases for one-item and three-item token options, respectively (two-tailed binomial test, $p < 0.05$). Note that Panzee only completed 15 trials of each type in the phase represented by section b.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript



a. Comparisons between 1 three-item token and multiple one-item tokens



b. Comparisons between 2 three-item tokens and multiple one-item tokens

Figure 4.

Chimpanzees' performance in Phase 3 of Experiment 1 as a function of the number of one-item and three-item tokens being compared. The x-axis represents the level of chance performance, so bars above and below the x-axis represent selection biases for the one-item token(s) and three-item token, respectively. Lower case o's and t's denote statistically significant biases for one-item and three-item token options, respectively (two-tailed binomial test, $p < 0.05$).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

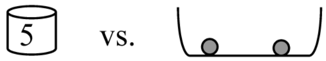
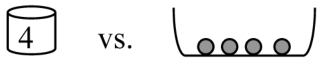
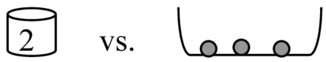
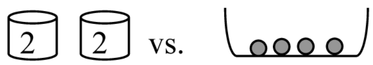
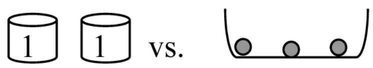
The Six Trial Types Involving All Food Containers (Experiment 1, Phase 2)

Trial Type	Description	Example Trial
Concordant cues	The correct set has a greater number of containers and higher value containers	4 4 vs. 2
Number cue	The value of the containers is the same, but the number of the containers is different	2 2 vs. 2
Value cue	The number of containers is the same in both sets, but their values differ	4 4 vs. 2 2
Discordant cues with number as the relevant cue	The correct set is in 2 containers of a smaller value rather than in a single container of a larger value	4 4 vs. 5
Discordant cues with value as the relevant cue	The correct set is in a single container of a larger value rather than in 2 containers of a smaller value	5 vs. 2 2
Equal	The number of items is equal in both sets, but the number and value of containers differed	4 vs. 2 2

Note: The number of depicted cylinders represents the number of containers present in a trial, and the cylinder labels represent the container values (i.e. the number of food items contained within).

Table 2

The Six Trial Types Involving Food Containers and Visible Food Items (Experiment 1, Phase 3)

Trial Type	Example Trial
One Container > Visible Set	
One Container = Visible Set	
One Container < Visible Set	
Two Containers > Visible Set	
Two Containers = Visible Set	
Two Containers < Visible Set	

Note: The left half of the Example Trial column represents the number and value of containers in the trial (as in Table 1). The right half of the same column represents the number of visible food items presented in the trial.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3

The Seven Trial Types Presented in Phase 3 of Experiment 2

Type	Source	Token Comparison
#1	Addessi et al. (2007)	3 vs. 1 1 1
#2	Addessi et al. (2007)	3 vs. 1 1 1 1
#3	Addessi et al. (2007)	3 vs. 1 1 1 1 1
#4	Addessi et al. (2007)	3 3 vs. 1 1 1 1
#5	Addessi et al. (2007)	3 3 vs. 1 1 1 1 1
#6	Addessi et al. (2007)	3 3 vs. 1 1 1 1 1 1
#7	New to the present study	3 3 vs. 1 1 1 1 1 1 1 1

Note: Gray cylinders represent aluminum tokens that were each worth 3 food items, and white cylinders represent plastic tokens that were each worth 1 food item.