Current Zoology 57 (4): 419–428, 2011

Comparing children's *Homo sapiens* and chimpanzees' *Pan troglodytes* quantity judgments of sequentially presented sets of items

Michael J. BERAN[1*], Julie S. JOHNSON-PYNN[2], Christopher READY[2]

[1] Language Research Center, Georgia State University, Atlanta, GA 30302, USA
[2] Department of Psychology, Berry College, Mount Berry, GA 30149, USA

Abstract We presented a quantity judgment task that involved comparing two sequentially presented sets of items to preschoolers and chimpanzees using nearly identical procedures that excluded verbal instructions to children. Trial difficulty in this task reflected the ratio difference between sets of discrete items where larger ratios (e.g., 0.80 as from comparing 4 to 5) were more difficult than smaller ones (e.g., 0.50 as from comparing 4 to 8). Children also completed verbal-based tasks probing the relationship between counting proficiency and performance on the quantity judgment task of sequentially presented identical sized items. Both species' performance was best when ratios between comparison sets were small regardless of set size in all types of tasks. Generally, chimpanzees and older children performed better than younger children except at larger ratios. Children's counting proficiency was not related to success in choosing the larger of two quantities of identical-sized items. These results indicate that chimpanzees and children share an approximate number sense that is reflected through analog magnitude estimation when comparing quantities [*Current Zoology* 57 (4): 419–428, 2011].

Keywords Numerical cognition, Quantity judgments, Children, Chimpanzees, *Pan troglodytes*

Making decisions about differing quantities is a skill that forms a basis of mathematical knowledge in children (e.g., Gelman and Gallistel, 1978; Fuson, 1988; Bideaud et al., 1992; Donlan, 1998; Mix et al., 2002; Cordes and Gelman, 2005). This ability also is shared with nonhuman animals (Boysen and Capaldi, 1993; Brannon and Roitman, 2003). It has become well-established that nonhuman animals can be very good at quantifying all kinds of things in their environment. They can make relative quantity judgments in which two sets are compared on the basis of the quantities within them (Addessi et al., 2008; Beran and Beran, 2004; Brannon and Terrace, 2000). These judgments are so widespread that they have been demonstrated in nearly all species tested to date, including great apes, monkeys, pigeons, dolphins, parrots, horses, dogs, voles, fish, and salamanders (Uller et al., 2003; Ferkin et al., 2005; Jaakkola et al., 2005; Emmerton and Renner, 2006; Hanus and Call, 2007; Dadda et al., 2009; Uller and Lewis, 2009). Some animals also can show counting-like abilities as they label or create sets to match a cardinal value (e.g., Matsuzawa, 1985; Boysen and Berntson, 1989; Pepperberg, 1994; Beran and Rumbaugh, 2001).

An emerging consensus is that quantification in the absence of a formal number system and absence of the use of a counting routine is mediated by a nonverbal/preverbal analog mechanism that uses mental magnitudes to represent quantity in an increasingly inexact manner as a function of true set size (e.g., Huntley-Fenner and Cannon, 2000; Huntley-Fenner, 2001; Brannon and Roitman, 2003). There also is consensus that humans and nonhuman animals share this analog system (Gallistel and Gelman, 2000; Cantlon and Brannon, 2006; Jordan and Brannon, 2006), although other mechanisms may play a role in the enumeration and quantification of small sets of items (e.g., object files, which allow for tracking small numbers of individual elements: Hauser et al., 2000; Feigenson et al., 2002). To date the evidence from studies with humans and nonhumans suggest that they form approximate representations of quantities, which may be dependent on task constructs (e.g., set sizes, continuous versus discrete quantities, sequential versus simultaneous and visual versus auditory presentation of items). In any case, the comparative approach will aid in discerning

Received Dec. 02, 2010; accepted Feb. 20, 2011.
* Corresponding author. E-mail: mjberan@yahoo.com

whether the nonverbal numerical representation of nonhuman primates and the numerical representation of young children emerge from similar mechanisms (Hauser et al., 2000; Jordan and Brannon, 2006). Here, chimpanzees are used as the nonhuman comparison species because they are often presented with nearly identical tests as those presented to children.

When one compares two quantities, and cannot use a process such as counting whereby discrete, exact labels can be applied and then compared, and performance nicely matches that predicted by Weber's Law. Weber's Law states that a constant level of discrimination of two quantities requires increases in the difference between those quantities proportional to the magnitude of those quantities. Specifically, the ratio measure of two sets (as indicated by dividing the smaller set size by the larger set size) takes into account the numerical distance between sets and the overall magnitude of sets, and it is more highly predictive of performance than either distance or magnitude alone (e.g., Beran, 2004). Higher ratios (e.g., 4 vs. 5, ratio = 0.80) lead to lower levels of correct responding compared to lower ratios (e.g., 2 vs. 7, ratio = 0.28). Many animal species show this pattern during quantity judgments in which performance is constrained by Weber's Law (Gallistel and Gelman, 2000; Brannon et al., 2006; Jordan and Brannon, 2006).

The focus in this paper is on tests with sequentially presented items. Beran (2001, 2004), and Beran and Beran (2004) reported that chimpanzees performed at high levels when they were shown two sets of food items that were placed, one item at a time, into opaque containers and then were allowed to select one of the containers. Across all set sizes, success was determined by the ratio between sets, not on the basis of the magnitude of the sets. Subsequent use of the one-by-one sequential presentation method with rhesus monkeys (Beran, 2007), capuchin monkeys (Beran et al., 2008; Evans et al., 2009), gorillas, bonobos, and orangutans (Hanus and Call, 2007), and even adult humans who were prevented from counting the arrays (Beran et al., 2006) indicated this same relation between performance and ratio between sets. This suggests that a similar mechanism may be operating across species to facilitate performance on this task. This type of experimental task has yet to be presented to preschool-aged children for direct comparison with nonhuman animal performance. The first aim of the paper was to expand the comparative assessment of quantification of sequentially presented sets among chimpanzees and human children.

The second aim was to examine the relation between the counting skills of children at various ages and performance during this type of quantity comparison. It has been hypothesized that formal counting competence is necessary for making such quantity comparisons, and this hypothesis has important implications for understanding the potential phylogenetic distribution of various quantity judgment skills. However, this hypothesis largely has been rejected, suggesting that such counting competence is not necessary. First, human children not yet capable of counting proficiency often perform well on other tests dealing with quantity judgments (Brannon and van de Walle; see also Mix, 1999). Second, although a number of nonverbal, nonhuman animal species successfully make quantity judgments as noted above, they have not mastered the counting routine. A more parsimonious hypothesis is that children show a developmental trend for increased performance with age in the quantity judgment and counting tasks but that there is not a clear relation between counting proficiency and performance on the quantity judgment task. To test this hypothesis we presented preschoolers and adult chimpanzees with combinations of sets containing 1 to 12 discrete items across a range of ratios. We predicted greater accuracy for set sizes with smaller ratios (e.g., 2 vs. 7) than those with larger ratios (e.g., 2 vs. 3) for both species. In addition, we predicted roughly comparable performance between children and chimpanzees, and we predicted that children's performance would improve with age. However, we did not predict a relation between counting skill and performance in the judgment task for children.

1 Materials and Method

1.1 Participants

All participants were recruited based on parental or guardian consent obtained through collaboration between their laboratory school and the Departments of Psychology and Teacher Education at Berry College. Participants included preschoolers from a class of younger students (4 males and 5 females) and from a class of older students (5 males and 6 females). The mean age for younger children was 44.11 months (range = 41–47 months), and the mean age for the older children was 53.45 months (range = 49–59 months). Preschool-aged children were chosen because, generally, this age group has not received intentional formal instruction in counting and arithmetic (Bisanz, Sherman, Rasmussen, and Ho, 2005).

The chimpanzees were Lana, a 32- year-old female; Sherman, a 29- year-old male, and Panzee, a 17- year-old female. The chimpanzees were housed at the Language Research Center of Georgia State University. These animals have extensive experimental histories that include participation in previous studies relating to numerical cognition (e.g., Beran and Rumbaugh, 2001; Beran and Beran, 2004). Although two of these animals (Lana and Sherman) had completed similar tests previously (Beran, 2001, 2004), the data presented in this experiment are from new test sessions with these animals and do not include any previously collected data.

1.2 Experimental Design

The quantity judgment task used here is virtually identical to that used with chimpanzees in previous studies (e.g., Beran, 2001, 2004), including two of the chimpanzees in this experiment. To more fully equate the two tasks, children were not instructed to attend to quantity. However, the task structure made it adaptive to do so. We avoided the use of instructions to determine if children, like chimpanzees, would choose the larger of two sets spontaneously, without verbal instructions from the experimenter. This methodological detail, of withholding instruction, sometimes leads to poorer performance in children compared to either nonhuman animals (Tomasello et al., 1993) or younger counterparts (Overman et al., 1996), but we expected that for a quantity judgment task, children might show a natural bias to "go for more" (Estes, 1976) in much the same way as has been demonstrated with nonhuman animals (Rumbaugh et al., 1987; Boysen and Berntson, 1995).

Children were tested individually in their classrooms and were *not* given instructions as to the object of the task (i.e., "choose the larger quantity"). Children were instructed to watch closely and to pay attention. Verbal praise for task participation (e.g., "You're doing a great job.") was given by the experimenters during testing but was not contingent on performance (i.e., participants were praised for being "on task" regardless of whether they chose the larger quantity).

During a trial, the participant and Experimenter 1 sat across from each other at a small table. Two opaque containers were placed on the table a sufficient distance away from the child so as to prevent viewing of the containers' contents. A clear plastic bag holding a quantity of identically colored beads was accessible to Experimenter 1. Experimenter 1 reached into the plastic bag and removed a quantity of beads, while keeping them hidden in his or her hand. The quantity was more than the number of items to be deposited into cups in order to eliminate any possible visual cues of the number of items in hand that might inadvertently occur. Next, Experimenter 1 dropped the predetermined number of beads into the container on his or her right, one at a time, using a quasi-randomized pace of placement in order to avoid rate of bead-dropping as a cue, (i.e., the larger quantity did not always fall at a faster rate). Experimenter 1 looked down while dropping the beads so that the child could not see the experimenter's face, thereby eliminating potential facial cues. The same procedure was followed to deposit items into the second container. Whether the larger quantity was deposited into the right or left container varied randomly from trial to trial (i.e., the larger quantity was not always presented in the right cup). Experimenter 1's right hand was always held over the right container for a longer length of time to control for temporal cues (i.e., the experimenter's hand did not always remain longest over the cup with the larger amount). These temporal controls meant that a child could not just use the length of time the hand was over a container or the consistency in the rate of object dropping to determine the larger set.

When finished, Experimenter 1 prompted the child and Experimenter 2 to proceed by stating, "Ready? Ok" while turning his or her gaze away from the test area. Experimenter 2, who was seated facing away from the child and Experimenter 1, did not know the number of beads in the containers. When hearing this oral signal, Experimenter 2 picked up the two containers, without viewing their contents, and presented them to the child while saying "Ok, you choose a cup." The containers were held in front of the child, so that their contents were not visible. After the child selected a container by touching it, Experimenter 2 poured out its contents on the table and announced to the child, "Ok, you chose this container, and you get to keep these beads." Next, Experimenter 2 poured out the contents of the container not chosen by the child. The participant was given the beads from the container he/she selected, while Experimenter 1 was given the beads from the remaining container. Although children did not receive any tangible reward until the end of the session, when their beads were returned in exchange for the chance to select a toy prize from among a number of alternatives, all children showed motivation to select containers.

A block of 17 magnitude comparisons was presented to children in a randomized order (see Table 1), and these comparisons included a range of ratios from small (0.20) to large (0.80). Nineteen children completed one block of trials in each of 6 test sessions for a total of 102 trials per child. One child completed 5 trials with each comparison (85 trials); this participant fell ill during the

Table 1 Characteristics of comparisons presented to children and chimpanzees

Sets to be compared	Ratio between sets	Interval distance	Absolute set size
1 vs. 5	0.20	4	6
1 vs. 4	0.25	3	5
2 vs. 7 *	0.285	5	9
3 vs. 10 *	0.30	7	13
1 vs. 3	0.33	2	4
2 vs. 5	0.40	3	7
1 vs. 2 *, 2 vs. 4 *, 4 vs. 8 *, 5 vs. 10 *	0.50	1, 2, 4, 5	3, 6, 12, 15
3 vs. 5, 6 vs. 10 *	0.60	2, 4	8, 16
2 vs. 3 *, 4 vs. 6 *, 8 vs. 12 *	0.667	1, 2, 4	5, 10, 20
3 vs. 4	0.75	1	7
4 vs. 5	0.80	1	9

Comparisons marked with an * are the ones given to the chimpanzees. All 17 comparisons were presented to children.

study and could not continue.

The test procedure used with chimpanzees is described in Beran (2001, 2004) and followed that used with the children. The only procedural difference was that the chimpanzees were given mini marshmallows from the container they selected, whereas the children traded their beads for a prize at the end of the day's test session. There were always two experimenters present for sessions with chimpanzees to control for inadvertent cuing of the animals.

Ten different magnitude comparisons were presented to chimpanzees in a randomized order during these sessions (see Table 1). Chimpanzees completed a total of 100 trials over four separate test days. Because a corpus of data from chimpanzees already indicated clearly the behavioral pattern that emerged from a wide variety of numerical comparisons on this task, we focused on a smaller number of comparisons here to provide representative data across the range of comparisons used with the children as well as specific data from novel comparisons never before used with chimpanzees in this task (e.g., 8 versus 12). In this way, we could evaluate the generality of quantity judgments across species with this task. The reader is directed to Beran (2001, 2004) for a larger corpus of data from this species using the identical task.

To assess enumeration and formal counting, children were asked two questions similar to the procedures of Mix et al. (1996), Rousselle et al. (2004), and Wynn (1990) at the end of the first, third, and sixth test sessions. First, the experimenter dumped beads in front of the child on the table workspace and asked, "Can you put 3 beads in this cup?" The same question was repeated for 6, 8 and 12 beads. The child was asked to collect a specific number of beads and was not instructed to count, but counting was permitted if children did so spontaneously. This provided data on cardinality (i.e., whether the final count corresponds to the count requested by the experimenter). Next, the experimenter put 3 beads in a pile in front of the child on the table workspace and asked, "Can you count these beads for me?" The same question was asked for 6, 8 and 11 beads. This provided additional data on counting procedures used by children (e.g., demonstration of one-to-one correspondence by finger tagging items, verbal errors of omission and repetition).

1.3 Test materials

Test materials used with child participants consisted of identically colored plastic beads (available at craft stores) and opaque containers. The opaque containers used with children (height = 9 cm diameter = 5 cm) were lined with foam in order to remove sound cues of beads being dropped into them. The same types of materials were used to test chimpanzees, except the cups were larger (height = 12 cm, diameter = 10 cm), and mini-marshmallows were used instead of beads.

1.4 Analysis

Children's performance in the quantity judgment task was scored for all 17 comparisons presented during each test session (test days 1–6), yielding mean percentage accuracy scores for each set comparison (across test days) and for each test session. A Repeated Measures ANOVA tested for learning over the six days of testing (unit of analysis was each child's mean accuracy score from each of the six test sessions). Children's performance did not differ significantly over the six test sessions (Ms for sessions 1-6, respectively = 62.3; 64.15; 65.1; 60; 60.5; 64.63, $F_{5, 75} = 0.99$, $P = 0.43$), thus, we combined scores from the six test sessions for subsequent analyses.

We adopted a hierarchical correlation approach to analyze whether aspects of trial composition influenced accuracy in choosing the larger of two sets. We performed a step-wise multiple regression to analyze the extent to which numerical scales of comparison sets (ratio, interval distance, and absolute set size) influenced performance (unit of analysis was each child's mean accuracy score for each comparison set from six test sessions). We predicted that ratio would be significantly related to accuracy, whereas interval distance and absolute set size would not. Consistent with other studies on children's numerical judgments, we predicted that older children would be more accurate than younger children in selecting the larger of two sets, but that even younger children's performance would be above chance, given that counting proficiency was not necessary for the quantity judgment task. A *t* test was used to determine if children's overall performance was significantly above chance.

Proficiency scores for collecting requested numbers of items (3, 8, 12) and counting presented sets (3, 6, 8, 11 items) were computed for each child after test sessions 1, 3, and 6. Responses were coded as correct if the child collected exactly the requested number of items or counted out the requested quantity with no mistakes. Our scoring criteria for each trial was more stringent than what some other research groups have adopted, however, we provided each child with three opportunities to count each quantity and averaged these to yield a single score for each quantity. This score did not indicate that the child had no knowledge or skill with the counting routine. Rather, it suggested that the child was not counting each set and comparing their cardinal values as the process for making quantity judgments in the magnitude comparison tasks. Children's performance in either task did not differ significantly over the three sampling periods (Repeated Measures ANOVAs for enumerating (WS variable: three, eight, and twelve, $P > 0.05$) and for counting (WS variable: three, six, eight, eleven, $P > 0.05$). Thus, the results reported for collecting and counting sets of beads were generated from data pooled across the three sampling periods. Correlation analyses compared scores for enumerating and counting sets to those on the magnitude comparison task.

Mean accuracy scores for five ratios (0.285, 0.30, 0.50, 0.60, 0.67) generated from comparisons that were presented to both children and chimpanzees were compared using a Mixed Design ANOVA, with ratio as the within subjects factor and species as the between subjects factor.

2 Results

Children performed reasonably well in the quantity judgment task (M = 63.12%, SD = 0.21, range = 48%–78%), and collectively, their accuracy in choosing the larger quantity was significantly above chance for all comparisons combined, t_{19}= 8.01, $P < 0.0001$ (two-tailed). They did this without instruction from the experimenter. Only one participant, a four year-old female, performed at chance levels (48%).

The chimpanzee data are compared to the data from all children in Fig. 1-Panel A. Children's combined performance was significantly more variable than that of chimpanzees at the 0.30 ratio ($P = 0.04$) and approached being significantly more variable at the 0.285 ratio (0.08); hence, the Greenhouse-Geisser correction was applied to the degrees of freedom. The three adult chimpanzees selected the larger of two sets with greater accuracy than children across the range of ratios, $F_{1, 21}$ = 9.84, $P = 0.01$, $n^2 = 0.32$, observed power =0.85, and both subject groups performed better at smaller ratios, $F_{2.8, 59.3} = 5.81$, $P = 0.002$, $n^2 = 0.22$, observed power = 0.93. There was no significant species by ratio interaction, $F_{2.8, 59.3} = 1.41$, $P = 0.24$. The observed power (0.35) was lower for this statistical comparison compared to the others.

A stepwise multiple regression analysis confirmed our prediction. The absolute size of comparison sets (pr = – 0.03, $P = 0.60$) and the interval distance between set sizes (pr = – 0.01, $P = 0.88$), did not explain significant amounts of variance in the model. Two variables contributed significantly to children's accuracy in choosing the larger set, $F_{2,339} = 11.44$, $P = 0.0001$. Preschoolers' percentage of correct choices was significantly impacted by the ratio between sets, $\beta = -0.22$, $P = 0.0001$, and their age (in months), $\beta = 0.13$, $P = 0.02$ (the increment in $R^2 = 0.02$ when age was added to the model), indicating that when ratios between sets were small, children were more accurate in choosing the larger quantity (see Fig. 1: Panel B), and that older children tended to choose correctly more often than younger children (see Fig. 2), although R^2 and confidence intervals for both variables indicated that the effects were small to moderate (ratio: lower bound = –0.37, upper bound = –0.13; age: lower bound = 0.01, upper bound = 0.009). Although narrower intervals indicate a better estimate than wider ones, the presence of zeros in the age variable confidence intervals might indicate the possibility of no

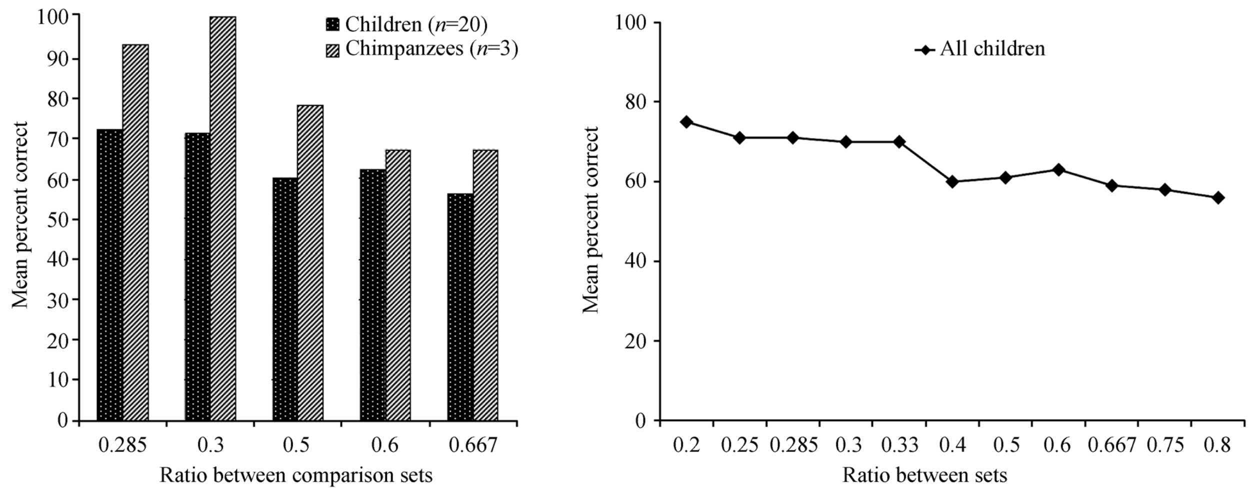

Fig. 1 Performance of the chimpanzees and children as a function of the ratio

A. Children and chimpanzees performed significantly better at smaller ratios compared to larger ones ($P < 0.01$, $n = 20$ for children, $n = 3$ for chimpanzees). **B**. A similar pattern was obtained when children were tested with other ratios.

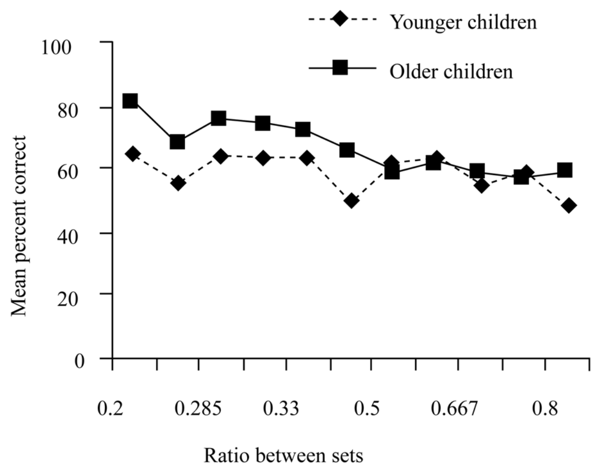

Fig. 2 Performance differences between age groups of children declined at larger ratios compared to smaller ratios

relationship in the population.

The age difference between older and younger preschoolers, however, diminished at higher ratios. At the six ratios that were less than 0.50, 45% of mean scores were in the 83%–100% range in children between 50–59 months of age, whereas only 21% of mean scores were in this range in children aged 41–49 months. For the five ratios varying from 0.50 to 0.80, the percentage of mean scores in the 83%–100% range was 14% for 50–59 month-olds and 9% for 41–49 month-olds. Only one child in the younger age group scored 100% correct (at 0.2 ratio), compared to 11 children in the older age group. Most of the perfect scores of older children were at lower ratios, and only one child had a 100% in a ratio above 0.50 (at 0.67).

Not surprisingly, children performed significantly better in collecting sets of 3 beads (M = 69%, SD = 0.46) than sets of 8 (M = 40%, SD = 0.49) and 12 (M = 17%, SD = 0.38), $F_{2,\,110} = 29.58$, $P = 0.0001$, $n^2 = 0.35$. There was a main effect of age; older children (M = 55.33, SD = 0.33) performed better than younger children (M = 25.67, SD = 0.42; $F_{1,\,55} = 14.30$, $P < 0.0001$, $n^2 = 0.21$). Only when collecting sets of 3 items did some children gather the exact number that had been requested. Post hoc dependent t tests were used when children erred to test whether their enumeration estimates for 3, 8, and 12 came close to the requested number. Non-significant t tests would indicate lack of a statistically significant difference between participants' responses and the quantities. This was the case for children who made errors with the quantity twelve, t_{19}= 1.56, $P = 0.14$, but not for eight, t_{17}= 2.68, p =0.02 or three, $t_7 = 4.20$, P = 0.004. Children who made errors enumerating twelve typically gave the Experimenter 13–14 beads (M difference = 1.51). Those children who erred enumerating eight gave the Experimenter 10–11 beads (M difference = 2.6), and those who made mistakes with three gave 9–10 beads (M difference = 6.83).

Children counted smaller sets (3: M = 92.5%, SD = 0.22; 6: M = 69.7%, SD = 0.25) with greater accuracy than they counted larger sets (8: M = 36.3%, SD = 0.31; 11: M = 30%, SD = 0.36), $F_{3,\,110} = 40.79$, $P = 0.0001$, n^2 = 0.43. Children from the older group (M = 64, SD = 0.39) performed better in counting overall than those from the younger group, although the magnitude of the difference was not substantial (M = 47.25, SD = 0.41, F (1, 55) = 6.42, $P < 0.05$, $n^2 = 0.11$). When asked to count beads, children typically employed a finger tagging strategy and often made errors of repetition and omission (i.e., they repeated and/or skipped numbers). For example, when asked to count 8 beads, one participant counted by finger tagging each bead while counting out loud, repeating the number "6" twice, skipping "7", and counting the last bead as 8.

As predicted, counting proficiency was not significantly correlated with performance on the magnitude comparison task for either counting (r = 0.12, P = 0.61, n = 20) or enumerating (r = 0.42, P = 0.07, n = 20; see Fig. 3). Given that children's performance differed depending on the quantity to be collected or counted, we conducted separate correlation analyses to rule out the possibility that performance on the quantity judgment task was related to performance on either the collecting or counting task. In other words, we wanted to compare children's best performance (at lower numerosities) in the counting tasks to their performance in the quantity judgment task so as not to underestimate their counting competency and its relation to numerousness judgments. None of these correlations were significant (P > 0.10 in all five cases).

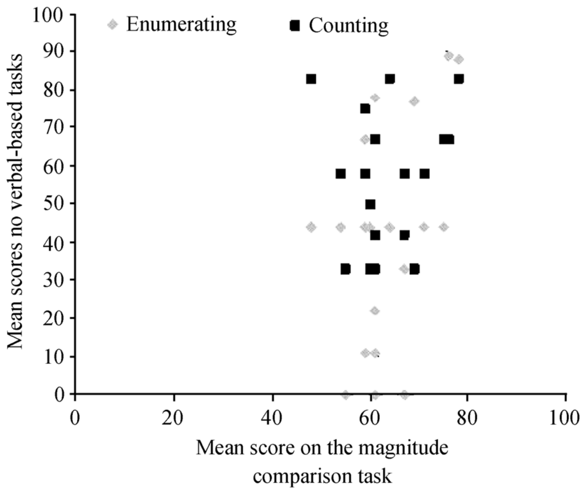

Fig. 3 Children's performance in the nonverbal magnitude comparison task did not depend on their ability to enumerate or count items in verbal-based tasks

3 Discussion

Children and chimpanzees were successful in selecting the larger of two successively presented quantities of items across a varied set of ratios. Children, like chimpanzees, attempted to choose the larger set without specific instructions. Their verbal comments during testing were indicative of their sensitivity to number and its relation to the sets they chose. After the experimenter finished depositing beads into the containers, children sometimes made remarks using number words (e.g., "You only have one"; "You have 5 and I have 10."), albeit erroneously (i.e., in most cases the number words only approximated the quantities). These results fit nicely with many previous studies that involved quantity judgments and showed clear ratio effects reflective of Weber's Law, including previous studies with chimpanzees (e.g., Beran, 2001, 2004; Beran and Beran, 2004). Thus, our hypothesis was confirmed that a comparative approach that included human children would show similarity across groups in performance. There is clear continuity in the performances of many species when making quantity judgments such as these. The present data support the increasingly clear consensus that an analog magnitude system exists for representing approximate quantities and numerosities, and this is an evolutionary widespread and ancient system.

The performance advantage the three chimpanzees showed over the children in this study could be attributed to the extensive test history of the chimpanzees, who often participate in quantity judgment tasks, or the differences in motivation prompted by test procedural differences. The chimpanzees worked for food, whereas the children worked for beads that were exchanged for toy prizes**.** Children also were rewarded at the end of the session, whereas chimpanzees were rewarded during the session after each trial. Nonetheless, the fact that children's verbal comments indicated both their recognition of the goal of the task (to pick the larger of two quantities) and their delight in picking the larger quantity when the contents of cups were revealed (e.g., by exclamations such as "I beat you.") lend validity to the comparison of the two species' performance. More importantly, their performance profile was the same.

The issue regarding the use of different to-be-enumerated items is very important. With chimpanzees, the use of highly preferred food items made it likely that the animals would attempt to maximize their gains from the very first trial (i.e., without training). However, we cannot say whether the animals were responding to the number of food items they saw or to the amount of food they saw. Because we did not know whether the chimpanzees used number or amount, we did not give the children verbal instructions in the form of telling them to pick on the basis of any magnitude. In other words, we did not want to use phrases such as "bigger amount," "more beads," or "larger set," because we did not want bias to control their responses in this task, thus making it less equivalent to the task as seen from the perspective of the chimpanzees. This manipulation offered us the opportunity to observe whether quantity controlled responding spontaneously as it does with nonhuman animals in situations in which one would expect them to use such information to maximize food intake (Call, 2000; Hauser et al., 2000; Beran, 2004). Ideally, we would have used food items with the children as well, but institutional guidelines at the testing location prevented us from doing so. Thus, although it was possible that children approached the task with lower motivation to select the larger set than did the chimpanzees this did not appear to affect children's motivation to participate in the task. They were attentive and engaged during testing and readily completed the trials, voicing their pleasure from getting beads from the experimenter.

Children, like chimpanzees, performed best when ratios between sets decreased, but performance was not affected by the total quantity of sets. Both subject groups were successful even when the absolute size of both sets combined was 20 items. Similar results were observed in Rousselle et al.'s (2004) study with preschoolers, whose magnitude comparison of sets of sticks showed a ratio effect but not a size effect, and in Hauser et al.'s (2003) study with tamarins, where ratio values determined the success of discriminating numerosities of tones. In addition, the data from the children nicely complement the data from rhesus monkeys (Beran, 2007; Cantlon and Brannon, 2006), adult humans (Beran et al., 2006), chimpanzees (Beran, 2001, 2004) and other apes (Hanus and Call, 2007) in showing the same general trend of ratio dependence in performance when comparing multiple, sequentially presented sets of items. Our results suggest that preschoolers and chimpanzees depended on an approximate numerical system, not an object-tracking system that detects one-to-one correspondence between limited number of objects presented visually and those stored in memory (e.g, Feigenson et al., 2002; Hauser and Carey, 2003; Hauser et al., 2000), or else their performance with small sets, such as the comparison 1 vs. 2, would have been much better.

Children's counting proficiency, however, was not related to their performance on the quantity judgment task, and this confirmed our hypothesis. Although children could not reliably count beyond six beads with complete accuracy, their performance in judging the larger of two sets containing more than six beads typically was above chance levels. This is suggestive of different cognitive processes being activated for comparing set sizes versus enumerating and counting individual items in sets. These results stand in contrast to those of Mix (1999) who reported that counting proficiency was related to performance in selecting the larger of two sets of items when each set was presented individually. However, that task and our task differ in that the Mix (1999) task involved matching a sequentially presented array to a picture with an equivalent number of dots (i.e., a set presented in its entirety). Therefore, children had to enumerate the sequential set and compare that information to a static set. This is more difficult than simply assessing the quantitative difference between two sets and choosing the larger, and thus may rely more on formal enumeration and counting skills. Our task is more like that used by Brannon and van de Walle (2001) that required comparisons of two sets of simultaneously presented items, and that produced data indicating that good performance was not related to mastery of the verbal counting system.

We attribute the success of children and chimpanzees in these quantity judgments tasks to a shared representational system, one based on estimation of relative magnitudes. In accordance with Weber's law, quantity judgments were more accurate when distance and magnitude of the differences between comparison sets was greater and was not negatively compromised by the number of items or size of the items in sets. Furthermore, precise judgments between comparison sets (e.g., 1 vs. 2) were not shown by children or chimpanzees, suggesting that they did not rely on an object-file based representational system.

Our finding that verbal counting knowledge was not related to children's performance in the quantity judgment tasks raises the possibility that the cognitive processes enabling formal counting and nonverbal quantity comparisons are not analogous. This view is consistent with the developmental progression of numerosity from estimation of relative magnitudes to mapping discrete quantities using a symbol system that supports formal counting and arithmetic and by neurological evidence that indicates dissociable neural systems for symbolic numerical competencies and nonverbal quantity representation (e.g., Nieder, 2009). Exact computation by addition or subtraction is by no means a characteristic of the average preschooler. Yet, preschoolers in this study showed some sensitivity, as did chimpanzees, to these very simple arithmetic operations. The ability to represent numerical transformations of this sort may be a prerequisite to the acquisition of precise mathematical skills.

Although we found no statistically significant relationship between preschoolers' performance in non-verbal quantity comparisons and formal-symbol based numerical tasks (even in older children and children who were competent counters), it is likely that pre-linguistic numerical competence is related to verbal counting in some way. Studies have demonstrated that minimal verbal numerical competence enhances children's performance in quantity mapping (Lipton and Spelke, 2005) and discrimination tasks (Brannon and Van de Walle, 2001). It may be the case that there is a bi-directional relationship between estimation and conventional counting because these two processes overlap in development as preschoolers learn verbal referents for quantities that had previously been represented non-verbally. However, the present data when examined from both the comparative perspective and as a function of comparing children's judgment performance with their counting performance suggests at least some independence between these two processes.

Acknowledgements We wish to acknowledge the staff at the Berry College Child Development Center and Georgia State University Language Research Center for their support in data collection. Research with chimpanzees was supported by grants from the National Institutes of Health (HD 38051 and HD 060563) and the National Science Foundation (BCS 0924811). The authors thank Mary Beran and John Kelley for their assistance in conducting experimental sessions with chimpanzees, and the staff at the Berry College Child Development Center for data collection with preschoolers. We thank Carla Moldavan for helpful comments pertaining to early childhood mathematical curricula.

References

Addessi E, Crescimbene L, Visalberghi E, 2008. Food and quantity token discrimination in capuchin monkeys *Cebus apella*. Anim. Cogn. 11: 275–282.

Beran MJ, 2001. Summation and numerousness judgments of sequentially presented sets of items by chimpanzees *Pan troglodytes*. J. Comp. Psychol. 115: 181–191.

Beran MJ, 2004. Chimpanzees *Pan troglodytes* respond to non-visible sets after one-by-one additions and removal of items. J. Comp. Psychol. 118: 25–36.

Beran MJ, 2007. Rhesus monkeys *Macaca mulatta* enumerate sequentially presented sets of items using analog numerical representations. J. Exp. Psychol. Anim. B 33: 42–54.

Beran MJ, Beran MM, 2004. Chimpanzees remember the results of one-by-one addition of food items to sets. Psychol. Sci. 15: 94–99.

Beran MJ, Evans TA, Leighty K, Harris EH, Rice D, 2008. Summation and quantity judgments of simultaneously and sequentially presented sets by capuchin monkeys *Cebus apella*. Am. J. Primatol. 70: 191–194.

Beran MJ, Rumbaugh DM, 2001. "Constructive" enumeration by chimpanzees *Pan troglodytes* on a computerized task. Anim. Cogn. 4: 81–89.

Beran MJ, Taglialatela LB, Flemming TM, James FM, Washburn DA, 2006. Nonverbal estimation during numerosity judgments by adult humans. Q. J. Exp. Psychol. 59: 2065–2082.

Bideaud J, Meljac C, Fischer J, 1992. The development of preschoolers' counting skills and principles. In: Bideaud J, Fischer JP, Greenbaum C, Meljac C ed. Pathways to Number: Children's Developing Numerical Abilities Hillsdale, NJ: Erlbaum, 99–126.

Bisanz J, Sherman JL, Rasmussen C, Ho E, 2005. Development of arithmetic skills and knowledge in preschool children. In: Campbell JID ed. Handbook of Mathematical Cognition. New York, NY: Psychology Press, 143–162.

Boysen ST, Bernston GG, 1989. Numerical competence in a chimpanzee *Pan troglodytes*. J. Comp. Psychol. 103: 23–31.

Boysen ST, Berntson GG, 1995. Responses to quantity: Perceptual versus cognitive mechanisms in chimpanzees *Pan troglodytes*. J. Exp. Psychol. Anim. B 21: 82–86.

Boysen ST, Capaldi EJ, 1993. The Development of Numerical Competence: Animal and Human Models. Hillsdale, NJ: Erlbaum.

Brannon EM, Cantlon JF, Terrace HS, 2006. The role of reference points in ordinal numerical comparisons by rhesus macaques *Macaca mulatta*. J. Exp. Psychol. Anim. B 32: 120–134.

Brannon EM, Roitman J, 2003. Nonverbal representations of time and number in non-human animals and human infants. In: Meck WH ed. Functional and Neural Mechanisms of Interval Timing. New York: CRC Press, 143–182.

Brannon EM, Terrace HS, 2000. Representation of the numerosities 1–9 by rhesus macaques *Macaca mulatta*. J. Exp. Psychol. Anim. B 26: 31–49.

Brannon EM, van de Walle G, 2001. The development of ordinal numerical competence in young children. Cognitive Psychol. 43: 53–81.

Call J, 2000. Estimation and operating on discrete quantities in orangutans *Pongo pymaeus*. J. Comp. Psychol. 114: 136–147.

Cantlon JF, Brannon EM, 2006. Shared system for ordering small and large numbers in monkeys and humans. Psychol. Sci. 17: 401–406.

Cordes S, Gelman R, 2005. The young numerical mind: When does it count? In: Campbell JID ed. Handbook of Mathematical Cognition. New York, NY: Psychology Press, 127–142.

Dadda M, Piffer L, Agrillo C, Bisazza A, 2009. Spontaneous number representation in mosquitofish. Cognition 112: 343–348.

Donlan C, 1988. The Development of Mathematical Skills. Hove, England: Psychology Press.

Emmerton J, Renner JC, 2006. Scalar effects in the visual discrimination of numerosity by pigeons. Learn. Behav. 34: 176–192.

Estes KW, 1976. Nonverbal discrimination of more and fewer elements by children. J. Exp. Child. Psychol. 21: 393–405.

Evans TA, Beran MJ, Harris EH, Rice D, 2009. Quantity judgments of sequentially presented food items by capuchin monkeys *Cebus apella*. Anim. Cogn. 12: 97–105.

Feigenson L, Carey S, Hauser M, 2002.The representations underlying infants' choice of more: Object files versus analog magnitudes. Psychol. Sci. 13: 150–156.

Ferkin MH, Pierce AA, Sealand RO, delBarco-Trillo J, 2005. Meadow voles *Microtus pennsylvanicus* can distinguish more over-marks from fewer over-marks. Anim. Cogn. 8: 182–189.

Fuson K, 1988. Children's Counting and Concepts of Number. New York: Springer Verlag.

Gallistel CR, Gelman R, 2000. Nonverbal numerical cognition: From reals to integers. Trends Cogn. Sci. 4: 59–65.

Gelman R, Gallistel CR, 1978. The Child's Understanding of Number. Cambridge, MA: Harvard University Press.

Hanus D, Call J, 2007. 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. J. Exp. Psychol. Anim. B 121: 241–249.

Hauser MD, Carey S, 2003. Spontaneous representations of small numbers of objects by rhesus macaques: Examinations of content and format. Cognitive Psychol. 47: 367–401.

Hauser MD, Carey S, Hauser LB, 2000. Spontaneous number representation in semi-free-ranging rhesus monkeys. P. Roy. Soc. Lon. B. Bio. 267: 829–833.

Huntley-Fenner G, 2001. Children's understanding of number is similar to adults' and rats': Numerical estimation by 5-7-year-olds. Cognition 78: 27–40.

Huntley-Fenner G, Cannon E, 2000. Preschoolers' magnitude comparisons are mediated by a preverbal analog mechanism. Psychol. Sci. 11: 147–152.

Jaakkola K, Fellner W, Erb L, Rodriguez M, Guarino E, 2005. Understanding of the concept of numerically "less" by bottlenose dolphins *Tursiops truncatus*. J. Comp. Psychol. 119: 286–303.

Jordan KE, Brannon EM, 2006. A common representational system governed by Weber's law: Nonverbal numerical similarity judgments in 6-year-olds and rhesus macaques. J. Exp. Child. Psychol. 95: 215–229.

Lipton JS, Spelke E, 2005. Preschool children's mapping of number words to nonsymbolic numerosities. Child. Dev. 76: 978–988.

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

Mix KS, 1999. Preschoolers recognition of numerical equivalence: Sequential sets. J. Exp. Child. Psychol. 74: 309–322.

Mix KS, Huttenlocher J, Levine SC, 1996. Do preschool children recognize auditory-visual numerical correspondences? Child. Dev. 67: 1592–1608.

Mix KS, Huttenlocher J, Levine SC, 2002. Quantitative Development in Infancy and Early Childhood. Oxford: Oxford University Press.

Nieder A, 2009. Prefrontal cortex and the evolution of symbolic reference. Curr. Opin. Neurobiol. 19: 99–108.

Overman W, Bachevalier J, Miller M, Moore K, 1996. Children's performance on "animal tests" of oddity: Implications for cognitive processes required for tests of oddity and delayed nonmatch to sample. J. Exp. Child. Psychol. 62: 223–242.

Pepperberg IM, 1994. Numerical competence in an African grey parrot *Psittacus erithacus*. J. Comp. Psychol. 108: 36–44.

Rousselle L, Palmers E, Noel MP, 2004. Magnitude comparison in preschoolers: What counts? Influence of perceptual variables. J. Exp. Child. Psychol. 87: 57–84.

Rumbaugh DM, Savage-Rumbaugh ES, Hegel MT, 1987. Summation in the chimpanzee *Pan troglodytes*. J. Exp. Psychol. Anim. B 13: 107–115.

Tomasello M, Savage-Rumbaugh ES, Kruger AC, 1993. Imitative learning of actions on objects by children, chimpanzees, and enculturated chimpanzees. Child. Dev. 64: 1688–1705.

Uller C, Jaeger R, Guidry G, Martin C, 2003. Salamanders *Plethodon cinereus* go for more: Rudiments of number in an amphibian. Anim. Cogn. 6: 105–112.

Uller C, Lewis J, 2009. Horses *Equus caballus* select the greater of two quantities in small numerical contrasts. Anim. Cogn. 12: 733–738.

Wynn K, 1990. Children's understanding of counting. Cognition 36: 155–193.