

Numerical Competence in Non-Human Primates: A Review of Indicators¹

Carla Krachun
Carleton U., Ottawa
ckrachun@chat.carleton.ca
May 2002

Introduction

Research on the numerical abilities of non-human primates centers largely around two questions: 1) do they have what might be called a number instinct? and 2) what is their number potential? The first question recognizes that number is a real property of groups of objects in the physical world. It is possible that animals could have evolved the ability to detect and respond to this property if it provided them with a selective advantage. This proposed evolved trait has been called many different things throughout the literature, including number instinct, number sense, number concept, numerical competence, numerical systems, numerical knowledge, and numerical representations, among others. In this paper, I use the term “number instinct” to mean an innate, hardwired ability which is either present at birth or which develops later as the organism develops and interacts with its environment, but is not the result of learning.

There is a large body of research suggesting that not just primates but also animals as distantly related to humans as rats and birds possess an innate responsiveness to the property of number (for a review see Dehaene, 1997). Human infants also display number sensitivities, even as young as a few days old (Wynn, 1998). Of course, none of these studies can prove that the creature in question had the displayed abilities from the moment of birth—they may have been acquired through early interactions with the environment. This does not take away from the fact that the *capacity* to acquire them rapidly through normal developmental experiences was obviously present, and this capacity itself could surely have been selected for in the species.

The second question above looks beyond innate abilities and asks to what extent can non-human primates be trained to use and understand numbers? If human infants and chimpanzees both display similar spontaneous abilities then perhaps chimpanzees can also develop more advanced numerical abilities with explicit instruction, just as human children do when they are taught to count and add. Two major undertakings of this sort are ongoing: Sarah Boysen’s experiments in the United States, and the work by Tetsuro Matsuzawa and his colleagues in Japan (reviewed in, respectively, Boysen & Hallberg, 2000; Biro & Matsuzawa, 2001). The abilities their chimpanzee subjects have purportedly developed through extensive training will be presented throughout the discussion that follows.

¹ Carleton University Cognitive Science Technical Report 2002-11
URL <http://www.carleton.ca/iis/TechReports>
© 2002 Carla Krachun

In this paper I will evaluate some aspects of the experimental research bearing on the questions above. I have compiled a list of indicators researchers use to infer what their subjects truly understand about the number tasks they participate in, or to infer what the underlying cognitive processes involved might be. These indicators are presented in turn, and examples from the literature are used to demonstrate how they have been employed. In some instances, where indicators overlap with those found in the non-human primate literature, examples are taken from the literature on humans and non-primate species.

Some of the indicators I present have served as formal dependent measures, others are features of the experimental tasks, while a few are simply informal observations that researchers have made in the process of carrying out experiments. As I proceed through the list, the conclusions researchers have drawn on the basis of each indicator are examined. The goal is to arrive at an overall sense of what the extant research can confidently tell us about the numerical instincts and potentials of non-human primates. Before plunging in I feel it necessary to address the “terminological chaos” (Davis & Perusse, 1988) that plagues this subject area, although I can unfortunately do little to quell it. My intention is to arm the reader with advance exposure to the issues so they will cause less confusion when they do arise.

Terminology

Subitizing

Kaufman, Lord, Reese, and Volkman (1949) observed that when adult humans were asked to label the number of items in an array there was a discontinuity in their responses: if the number of items was 6 or fewer the subjects performed the task very quickly, but beyond this number their response times increased at a steady rate along with the quantity of items. The authors decided the subjects were obviously counting the larger arrays, but what process were they using for the smaller ones? They invented the term subitizing, formally defining it as “the rapid labeling of small quantities of simultaneously presented items” (p. 520). This definition, however, is merely descriptive and reveals little about the underlying mechanisms at work.

Others since Kaufman have attempted to include in the definition of subitizing some explanation of the mechanism. Von Glaserfeld (cited in Davis & Perusse) emphasized that subitizing was a perceptual process rather than a cognitive or enumerative one, involving some form of pattern recognition using a flexible template. Hauser (1997) also describes subitizing as a low-level perceptual mechanism that happens “preattentively”, and Thomas, Phillips, and Young (1999) prefer to dispense with the term subitizing altogether and use prototype matching instead. Gallistel and Gelman (1992) are even more specific, linking subitizing to the theory that numerosities are represented internally as analogue magnitudes through an “accumulator” mechanism (first described by Meck and Church in 1983). They define subitizing as “the use of the preverbal counting process and the mapping from the resulting magnitudes to number words in order to generate rapidly the number words for small numerosities.” (p. 43)

Gallistel and Gelman’s (1992) definition brings up another point of contention among researchers: when is it appropriate to apply the term subitizing? Kaufman et al. (1949)

originally coined the term to refer to a process that occurred in human *adults* when they identified small numbers of objects. Gallistel and Gelman's definition also implies that the adult ability to use numerical symbols must be involved. Others do not require the process to involve the application of an external symbol as its end result, and are comfortable in also applying the term to what preverbal humans and animals do when perceiving small numerosities.

Counting

The terminological situation is no better for the word "counting", which has a variety of meanings throughout the literature. Davis and Perusse (1988) argue that the term should be restricted to formal enumeration in the adult human sense, which is the commonly understood meaning of the word. They characterize the difference between subitizing and counting as a "quantum leap" (p. 565), requiring abilities that may be present only in adult humans, in particular language. If this is the case then looking for counting in infants or animals may be misguided.

In contrast Gallistel and Gelman (1992) do not consider counting to be a process dependent on language, and so it *could* be within the behavioral repertoire of nonverbal creatures. They identify five principles of counting, which at first glance appear to depend upon the ability to use external symbols: one-to-one correspondence, stable order, cardinality, abstraction, and order irrelevance.

In truth, it is *not* possible to meet any of these criteria without using symbols, but Gallistel and Gelman suggest the symbols need not be external. Instead, they can be mental symbols, which they call "numerosons", internal tags the mind makes use of to enumerate a set of objects, perhaps something akin to bins in short term memory.

Estimation

Kaufman et al. (1949) suggest that estimation is really the same thing as subitizing, except that it happens when perceiving higher numerosities under time limits and is much more error-prone. In their experiments human adults appeared to stop subitizing and begin estimating when the arrays exceeded 6 items, unless they were given enough time to explicitly count the items. If time was limited, the number of mistakes they made increased sharply while their self-reported confidence in their judgments plummeted.

Gallistel and Gelman (1992) espouse a more inclusive definition in which "estimator processes" include any process in which real-world numerosities are mapped onto their numerosons. These are distinguished from "operator processes", which involve the manipulation of numerosons according to arithmetical rules to arrive at new numerosons, for example in addition. This distinction is important to make, but using the idea of estimation in a way that includes both subitizing and counting only further muddies the terminological waters. And to make things muddier still, there are those who believe that estimation is actually the most complex of the three processes, and the last one to emerge during human development (e.g., Klahr & Wallace, 1973, cited in Davis & Perusse, 1988). One cannot make meaningful numerical estimates, they argue, without already having a well-developed concept of number that includes the ability to count.

A final note on terminology

In the interests of brevity, the state of terminological affairs has been made above to appear less desperate than it actually is. There is much more I could have included, such as Davis and Perusse's (1988) proposed division of subitizing into the pre-counting and post-counting varieties; or the debate over whether subitizing can apply to items presented sequentially or just simultaneously; or the question about whether subitizing is possible in different sensory modalities; or the disagreement over whether estimation is a simpler or more complex process than counting; or where the terms "protocounting", "preverbal counting" and "relative numerosness judgments" fit into the picture. Some of these issues will come up in the discussion that follows and will be dealt with as needed.

It is important, however, to be clear on how the term "numerosity" is being used in the present paper. As distinguished from "numerals", which are verbal symbols, "numerosity" refers to the total collection of distinct units in an array. When I say that an animal was presented with a numerosity I mean that it was shown an array of items. When I say it was presented with a numeral I mean that it was presented with an Arabic number symbol. It is obvious that numerals can be mapped onto numerosities and vice versa, but they are not precisely the same thing.

The indicators

Indicators 1, 2 and 3: Speed and accuracy of response, and size of the array

In combination, these three variables are often used to decide whether the subject has subitized, estimated, or counted an array of items. This is made possible by the facts that subitizing and estimating are believed to be faster than counting, whereas both counting and subitizing are considered more accurate than estimating. In addition, subitizing is believed to be operant when the items in the array are few, while larger arrays must be either counted or estimated. Given these features of the three processes, if a subject is observed to judge the numerosity of a large array with slow speed but high accuracy then counting is assumed to have occurred. If the judgment of the same array is done very rapidly but inaccurately the inferred process is estimation. Whenever the array contains just a few items, allowing for a very quick and accurate judgment of its number to be made, then subitizing is assumed.

Thomas and Chase (1980) ruled out subitizing in their squirrel monkey subjects based on the fact that the subjects accurately distinguished numerosities of 7 from 8, and these numbers were larger than the maximum of 6 put forward by Kaufman et al. (1949). But this puts an unnecessarily stringent restriction on the size of array that can be subitized. The numerosity at which speed sharply decreases and subjects appear to begin counting (if they have time) or estimating (if time is limited) has *not* been consistently around 6 items for neither human nor non-human primates. The discontinuity in responses has often been observed around 4 items, and Thomas et al. (1999) found high accuracy in adults rapidly discriminating numerosities of between 10 and 11 items.

It is interesting to note that Thomas and Chase (1980) did *not* consider that their monkeys could have been counting instead of subitizing, because the subjects "lacked the requisite

training”. It was later proposed that the monkeys may have performed some kind of prototype matching that involved building prototypes from experience with different stimuli and then identifying novel stimuli by how closely they corresponded to those prototypes (discussed in Thomas et al., 1999). But this explanation is *not* inconsistent with subitizing; it simply attempts to specify the underlying process by which subitizing might occur.

Hauser (1997) argues that results obtained by Rumbaugh and Washburn (1993) could not be accounted for by subitizing because the number of objects presented far exceeded the demonstrated limit for adult humans, “and there is no a priori reason to think that chimpanzees should surpass this limit” (p. 113). Of course, there is also no a priori reason why they should not. Some research has, in fact, provided support for greater subitizing ranges in animals than in humans (for example Pepperberg, 1987; Terrell & Thomas, 1990). Davis and Perusse (1988), however, have suggested the latter results may have had more to do with task features than with different subitizing abilities between the species.

Thomas et al. (1999) hypothesized that human subjects would use some process other than counting to judge which one of two arrays had more dots, and also to judge which of two polygons had the greater number of sides. Because they hypothesized a non-counting process, the researchers predicted no significant increase in response times as the number of dots or sides increased. When they in fact found no difference in response speed across conditions they cited this as evidence that the subjects were not counting.

The problem with arguments such as that above based on speed, number, or size of array is that they have an unsettling circularity. Subitizing is defined as quick and accurate numerical labeling of small arrays, and so if a small array is labeled quickly and accurately it must have been subitized. If the array was large, however, it must have been counted because, by definition, subitizing does not occur with large numbers. If the array was large and labeled *inaccurately* it was probably estimated, because estimation typically results in inaccurate labels while counting does not.

It would be interesting to try and interpret the rapid and accurate labeling of a large array, or the slow and inaccurate labeling of a small array. By the logic of many of the researchers cited such things would be inexplicable. And, ironically, the best estimators—those who get it right almost every time—would never be identified as estimators precisely *because of* their accuracy. There are no doubt situations in which speed, accuracy or size of array, in conjunction with other indicators, may provide some clues about inner processes. But researchers should not allow proposed definitions to be a hindrance rather than an aid toward gaining a deeper understanding of underlying processes.

Indicator 4: Response pattern

Closely related to the use of response speed (discussed above) to infer underlying processes is the analysis of response patterns across conditions. Kaufman et al. (1949) originally determined that subitizing was a different process from estimation because of a sudden and obvious discontinuity in the slope of response speed after the array exceeded 6 items. At or below that number the slope was nearly flat, with each successive increment in numerosity being accompanied by a negligible increase in time taken to respond. Beyond 6 items, however, the slope became much steeper, such that each increment in numerosity was now accompanied by a much larger change in response time. Researchers use deviations from this typical human response pattern in animals as a strong indication of differences in the animal's underlying thought processes.

Biro and Matsuzawa (2001) are confounded by the fact that their chimpanzee's response times for labeling 1 to 9 items follow an unusual pattern. Ai's response time was the same for arrays of 1, 2 or 3 items; and it thereafter increased monotonically for each additional increment in the size of the array just as it does for humans, with one exception. For the array with the largest number of items that Ai had been trained to label, her response time declined. So, for example, when the extent of her training was with arrays of between 1 and 7 items, Ai's response time was flat for 1, 2 and 3, it increased monotonically for 4, 5 and 6, and then dipped again for 7 items. Later, when she learned to also label arrays containing 8 items, it was the 8-item array that experienced the dip rather than the 7. The same thing happened again when arrays of 9 items were introduced; Ai's response time stayed flat for 1 to 3 items, increased monotonically from 4 to 8 items, and then decreased for 9 items.

The investigators suggest subitizing as the possible mechanism used by Ai for the lower numerosities, but they are at a loss to explain the pattern after that point. It could indicate some kind of sequential tagging or "counting-like" behavior, they speculate, but the drop in latency for the highest number is difficult to reconcile with this explanation because this is not the pattern typically displayed by humans when they count. They conclude there must be some other process at work, but they cannot immediately identify what it might be. Giambrone (personal communication, 2002) suggests one alternative possibility. If Ai is engaging in some sort of perceptual pattern recognition then arrays with neighboring numerosities might be difficult to distinguish from one another. When Ai is viewing an array of 6 items, there will be two neighboring numerosities for her to contend with: 5 and 7. When she is viewing an array of 9 items, however, she needs to deal with only one other competing numerosity: 8. This may save her some deliberation, resulting in the observed dip in her response time. This is just speculation, of course, but it provides a starting point for thinking about how Ai's unusual response pattern might be explained.

Indicators 5 and 6: Self-reports and confidence ratings

The easiest way to determine whether subjects are subitizing, estimating or counting may appear obvious: just ask them. Thomas et al. (1999) gave their subjects a post-experiment questionnaire asking them to state whether they had counted during the number-identification tasks. Eighty percent of the subjects answered no, and the remaining 20% were unable to describe just *how* they made their judgments. This points to a common difficulty in using introspection: thought processes are often not available on a conscious level. Another obvious problem is that when the subjects are preverbal humans or animals they simply cannot self-report on their inner experiences.

Confidence in choice (in tasks in which the subject must indicate the greater of two arrays, for example) is another indicator made use of by some researchers, as higher confidence in one's choice is presumed to be associated with counting or subitizing rather than estimating. Kaufman et al. (1949), for example, measured subjects' confidence in their judgments of the number of dots on a screen by asking them to rate on a 5-point scale the accuracy of each of their reports, with 5 being absolute certainty and 1 being absolute uncertainty. They also cite a study by Taves (1941) in which confidence ratings were used to assist in making inferences about whether the subjects were counting or estimating. It is not possible to get direct confidence ratings from animals, of course. But their behaviors while completing a task, for example vacillating between two choices, might provide an indirect measure of confidence. Such clues might best be classed as "behavioral indicators", discussed below.

Indicator 7: Behavioral indicators

Behaviors that subjects engage in while carrying out numerical tasks have been used to argue for counting. Boysen, Berntson, Shreyer and Hannan, (1995) describe how Sheba began to spontaneously employ "motor tagging" or "indicating acts" —touching, moving, and rearranging the items in the array—before choosing the correct Arabic numeral. But it is possible that Sheba's manipulation of the items was simply a function of her inability to inhibit her interest in them, especially given that they were food items. This is supported by the observation that in a later study by Boysen, Berntson, Hannan and Cacioppo (1996), Sheba and another chimpanzee could not inhibit their tendencies to select the larger of two candy piles, even though a reversed reinforcement contingency dictated that choosing the larger pile resulted in gaining access to the smaller pile and vice versa.

Suzuki and Kobayashi (2000) described tentativeness in the behavior of their rat subjects as they passed each tunnel in a sequential array of tunnels, searching for the one that contained a food reward. This might indicate counting, but it could just as easily be based on the discrimination of a rhythmic pattern produced by the rats' own movements along the sequence of tunnels (Davis & Perusse, 1988). And although the experimenters did their best to rule out positional, tactile, visual and olfactory cues, it is possible the animals were picking up on variables not considered and therefore undetected by the experimenters. In fact, the rats' "stopping to check" behavior is highly consistent with such an interpretation.

Biro and Matsuzawa (2001) describe a couple of behavioral indicators they used to infer Ai's thought processes. The first occurred in a task in which an array of dots was presented to Ai along with a collection of Arabic numerals. Before choosing a numeral, Ai would often look from the numerals back to the array of dots, as if to be sure she was making the right choice. If Ai was actually counting the dots there would be no need to glance back and forth between the array and the numerals as she did, the experimenters reasoned, and so she must have been using a form of estimation (by which they appear to mean some type of perceptual recognition of the stimulus configuration).

The second behavioral indicator referred to by Biro and Matsuzawa (2001) occurred while Ai was ordering arrays of numerals in ascending sequence. The researchers noticed that her response latency was longest for the first numeral in the sequence, and was flat for all the other numerals (i.e., an L-shaped function). They hypothesized that Ai was pre-planning the motor actions required to touch all the numerals in the proper sequence, even before she made her first move. To test this they waited until Ai had indicated the first numeral in the sequence, and then quickly switched the locations of the second and third numerals (using a computer allowed them to do this without disturbing their subject). Not only did Ai's accuracy sharply decline when this was done, but her behavior strongly indicated the switch interfered with a pre-planned motor sequence. After touching the first numeral in the sequence, Ai's finger continued toward the original location of the second numeral, although the third numeral now occupied that location. In some instances Ai noticed the change before she had touched the incorrect numeral, and when this happened she changed course midstream and redirected her finger to the new location of the second numeral.

Tomonaga and Matsuzawa (2000) found evidence of pre-planning as well when they observed that Ai's latency to begin an ordering task became longer as the number of items to be ordered became greater. Kawai (2001) made a similar observation, and also found that when arrays included repetitions of some numerals (e.g., 2, 4, 3, 1, 3) Ai's time to begin increased along with the number of repeated numerals. This was probably, they conjectured, because Ai had to consider more than one possible path for touching all the items in sequence. And the more repeated numerals there were in the array, the more possible pathways she could take.

While these studies may provide some insight into how Ai carried out her ordering task (i.e., pre-planned the motor sequence and then carried it out), they do not offer much insight into how well she understood what it meant for the numbers to have an ordinal sequence. She could simply have learned to produce a particular sequence of symbols to receive a reward. The fact that she planned the motor actions in advance required to produce the proper sequence says nothing about her numerical understanding as such.

Indicator 8: Stimulus exposure time

The length of the stimulus exposure has also been used to infer, and in some cases even to induce, particular enumerative processes. To get humans to estimate, rather than count, the items in an array, the array may be presented for a very brief period, usually a fraction of a second. Of course, if the subjects are still allowed to take their time in providing their estimate they may hold a visual image of the array in their minds and use it to perform “mental counting”. To avoid this the experimenter might also ask subjects to provide their responses within a given time limit.

Biro and Matsuzawa (2001) used limited stimulus exposure times to investigate the reason behind Ai’s unusual response curve in a number labeling task (i.e., the final dip in response latency to the largest numerosity, described above under ‘Indicator 4’). In an attempt to explain the dip, the researchers gave Ai and also some human subjects the labeling task again, but this time they limited the stimulus presentation to 100 milliseconds. The humans showed the same steady monotonic increase they did before, even with the brief stimulus (the authors attributed this to mental counting). Ai, on the other hand, now showed a flat response latency function for *all* numerosities, which was explained by suggesting she must have resorted to estimation.

The researchers unfortunately did not report accuracy levels, which would have also been revealing. Were Ai’s estimates more often incorrect than the humans’ mental counting calculations? If she really was estimating we might guess they were because, as discussed previously, estimation tends to be less accurate than counting. If Ai was not estimating then what *was* she doing? Could she simply have counted much faster when she perceived the exposure time would be limited? This is less probable than the estimation account, but without further indicators to help infer the processes at work all possibilities must be considered.

Indicator 9: Logical transitivity and ordinality

In a much-cited report, Brannon and Terrace (1998) present evidence that rhesus monkeys represent numerosities from 1 to 9 on an ordinal scale. Most impressive is that the monkeys were able to order stimulus exemplars containing between 5 and 9 items, even though the animals’ training had only included exemplars containing 1 to 4 items. In other words, having learned to put exemplars of between 1 and 4 items in ascending order, the monkeys were then able to accurately order all possible pairs of numerosities between 1 and 9. This was the case even when both of the exemplars in the pair contained numerosities the animal had never before encountered (for example, 5 and 9, or 8 and 6). The results are intriguing, and may in fact demonstrate the authors’ claim that “monkeys can spontaneously represent the numerosity of novel visual stimuli and...can extrapolate an ordinal rule to novel numerosities” (p. 748). As the authors point out, however, whether or not the monkeys are succeeding in the task by employing processes that can be described as “counting” remains to be determined.

Indicator 10: Transfer

The chimpanzee Ai's early training involved mapping numerals to groups of everyday objects, such as toothbrushes, spoons and gloves (Biro & Matsuzawa, 2001). When the stimuli were changed to green dots positioned randomly on a computer monitor Ai's skills did *not* readily transfer, and she only learned the task after more than 150 training sessions. If Ai truly understood the concept of number based on her training with the everyday objects, should it have taken her so long to learn the dot labeling? Eventually, the researchers tell us, Ai's performance on the "dot -to-number matching test" (DNMAT) was highly accurate, but even they concede this does not imply she was counting.

Ai was next tested to see if she could transfer her newly acquired knowledge on the DNMAT to the reverse task, matching numerals back to dots (NDMAT). Once again, ready transfer between the tasks was *not* observed, revealing, in the researchers own words, "numerals were not used as 'symbols' in the strictest sense by the chimpanzee" (p. 212). After Ai had been trained to a high level of success on the NDMAT, she was next tested with adjacent pairs of numerals (1-2, 3-4, etc.) to see whether she could touch them in ascending order. There was no transfer, revealing she did not understand the ordinal relationships. And it was not just that it took Ai some time to catch on to the task requirements; after learning to order 1 and 2 she still had to be explicitly trained on 3 and 4, and then on 5 and 6, and again on 7 and 8. Moreover, when new adjacent pairs were tested (2-3, 4-5, 6-7), Ai performed *worse* than chance, responding on the basis of her prior reinforcement history rather than on the basis of any ordinal understanding. She picked 3 first because that is what she had been rewarded for in the past, for example. Ai was obviously good at learning contingencies, but she had no apparent understanding of the concept of ordinality.

Still, Ai was trained to order all possible pairs of numbers, and when she was thereafter given a novel pairing (8-9) she performed with high success (87.5%). Biro and Matsuzawa (2001) interpret this as evidence that Ai had finally grasped the ordinal relationships between the numbers, but there is a simpler explanation. Upon first seeing the novel pair, Ai could have avoided choosing 9 first because she was unfamiliar with it in the context of that particular task. That leaves 8 as the only other choice, so she chose 8 first and then selected 9 when there were no other options left. It might only take one or two trials of this sort before Ai, who is obviously very skilled at recognizing contingencies, would learn to consistently pick 8 before 9 in order to obtain a reward. That being said, it is impressive that Ai was able to later transfer her ordering abilities to non-adjacent pairs on which she had never been trained, such as 2-4, or 7-5. It is difficult to explain this without accepting that Ai understood the ordinal nature of the numbers.

It should be noted here that Sarah Boysen reported Sheba had little trouble transferring her ability to label numerosities to novel items (Boysen & Hallberg, 2000). And Sheba was apparently able to easily acquire "receptive comprehension skills" (equivalent to Ai's NDMAT skills) based on her previous learning with productive tasks (equivalent to DNMAT). In contrast to the Japanese researchers, who are conservative in their claims about Ai's true understanding of numbers, Boysen is highly confident in claiming

throughout her research that Sheba has a highly developed number concept. Future replications of the research may help to resolve the inconsistency, if anyone is ever again willing to take on the lifetime commitment of trying to teach mathematics to a chimpanzee!

Indicator 11: Types of error made

If there was a prize for pointing out the errors your subjects make and trying to learn from them, Biro and Matsuzawa (2001) would win that prize. Throughout two decades of training their chimpanzee Ai has made many mistakes and she continues to make them, especially when it comes to the quantities zero and one. She often labels zero items as 1, selects one item when presented with the numeral 0, selects zero items when presented with the numeral 1, and puts 1 before 0 in sequential ordering tasks. The researchers admit these persistent errors reveal Ai has little understanding of the role of 0 as a label for the absence of items. Further, they suggest she may be using relative magnitude judgments rather than counting to order numerals, because she has little trouble ordering 0 in relation to numbers larger than 1. Still, the difference between 0 and 1 is no less than it is between other adjacent numbers and Ai can order those without much problem, so the trouble may lie elsewhere.

Biro and Matsuzawa (2001) also speculate that Ai's difficulties with zero could be an artifact of her training history. The numbers 1 to 9 were taught to Ai in their proper sequence, with the next number in the sequence being introduced only after Ai had become proficient with the previous one. But zero was different. When it was introduced it was smaller than any of the other numbers previously learned, and it was placed at the beginning of the sequence rather than the end. The researchers conclude "the chimpanzee's competence in numerical tasks is a function not only of an underlying understanding of abstract numerical concepts, but also of the training history of the subject" (p. 218).

Biro and Matsuzawa (2001) also describe an experiment in which they varied the size and density of dots in the array to see whether these variables had an effect on Ai's performance in judging numerosities. Ai made slightly more errors as the proportion of large to small or medium dots in the array increased. Specifically, when the proportion of large dots was higher she picked a larger numerical tag. The researchers conclude this indicates a role for dot density, and not just dot numerosity, in Ai's judgments, making it clear that other perceptual variables must be carefully controlled for in experiments of this kind.

Indicator 12: Violation of expectancy and looking time

Researchers interested in studying the spontaneous numerical abilities of experimentally naïve animals have adopted some useful measures from studies with human infants. One of the advantages of taking this approach, they argue, is that it allows one to make direct comparisons between animals and preverbal humans. These comparisons may provide more insight into the evolution of human capacities than comparisons between humans and trained animals because they focus on biologically endowed, rather than learned, abilities (Uller, Hauser, & Carey, 2001).

The ‘violation of expectancy’ looking-time methodology is one example of an approach initially taken with infants and later adapted for use with non-human primates (others are detailed in Uller et al., 2001). Animals, like infants, tend to spend more time looking at a display when something happens that violates their expectancies or ‘surprises’ them. In one study Hauser, MacNeilage and Ware (1996) studied the responses of wild rhesus monkeys to changes in the number of eggplants on a display structure. While the monkeys observed, the experimenters placed either one or two eggplants behind an opaque barrier on the structure. In some cases, the experimenters then surreptitiously removed or added another eggplant and noted the subjects’ reactions when the barrier was removed. As predicted, the monkeys looked longer at the display when the number of eggplants did not match the number they had seen placed behind the screen. The experimenters suggested the monkeys were able to detect the results of addition and subtraction operations, providing support for the view that the animals possess some kind of basic numerical sensitivity or understanding. The researchers’ interpretation is cautious, however, and they also present the alternative explanation that the monkeys were using perceptually rich, high-level representations rather than innate arithmetical ability.

Uller, Hauser, and Carey (2001) found similar results with cotton-top tamarins, a New World species more distantly related to humans than rhesus monkeys. The animals looked longer at $1+1=1$ items than they did at $1+1=2$ items. And by comparing the animals’ responses to $1+1=2$ items versus $1+1=3$ items, the researchers determined it was specifically *two* items the animals were expecting to see, not simply ‘more than one’. They also eliminated the possibility that monkeys were responding to total mass rather than numerosity by demonstrating that monkeys looked longer at $1+1=1$ double-sized item than at $1+1=2$ items. The same amount of ‘stuff’ was there in both cases, but the monkeys appeared surprised when the actual number of items was not as expected.

The researchers in the above study claim their results provide evidence that tamarins can detect and respond to numerosities, but they concede the results say nothing about what kinds of representations the monkeys might use to do this. Nevertheless, it is clear from these and other studies using the same methodology (e.g., Sulkowski & Hauser, 2001) that monkeys show a robust ability to detect violations of simple mathematical rules, even when other possible factors are ruled out. We might conclude that, at the least, looking time as an indicator is useful in demonstrating sensitivity to numerosity in untrained non-human primate species.

Indicator 13: Preferential choice

Animals of all kinds are easily trained to choose between two stimuli to receive a reward, and the choices they make can indicate something about their numerical understanding. Presented with two trays, each containing several piles of candy, the chimpanzee Sheba consistently chose the tray with the larger total number, suggesting she was somehow adding the numerosities of the separate piles (Boysen & Berntson, 1995). Similar results were found with two chimpanzee subjects in experiments by Rumbaugh, Savage-Rumbaugh, and Hegel (1987).

Olthof, Iden and Roberts (1997) used the preferential choice paradigm with two squirrel monkeys, but in their study the animals chose among collections of Arabic numerals, rather than numerosities, to earn that same number of peanuts as a reward. Not only did the animals learn to choose the largest numeral, but when required to choose between *sets* of two numerals (e.g., 1,3 vs. 4,2) they showed a significant tendency to choose the set with the larger total number. This was true from the very beginning of their exposure to the sets, the authors point out, so it could not have been a result of *learning* the quantity of peanuts associated with each set.

What is most interesting about this study is that the researchers looked closely at the subjects' choices given particular pairings of numerals, in order to gain a more complete understanding of the monkeys' abilities. They noted, for example, that the monkeys were able to choose sums that exceeded any of the Arabic numerals they had experienced during training. So, for example, while the monkeys had learned that the symbol '5' earned them fewer peanuts than the symbol '7', they had not been exposed to the numeral 12 during training. Nevertheless, they appeared to realize that the symbols '5' + '7' together as a pair would earn them more peanuts than the symbols '3' + '3'. Of their results, the authors are careful to point out: 'nothing in these data suggests that the monkeys knew summation relationships among symbols or had an addition table in their heads' (p. 338). They offer a few possible reasons for the monkeys' success, including conditioned reinforcement as well as more cognitively based explanations. The monkeys' choices unfortunately could not reveal which of these explanations was most probable, but the study does make it clear that preferential choice can be very useful in at least determining the extent of a species' numerical abilities.

Conclusion

In this paper I have presented a variety of indicators researchers have used to investigate the innate and acquired numerical abilities of non-human primates. Some have been more useful to researchers than others, but there are limits on what any given indicator can tell us. Indicators of spontaneous number abilities in animals are indicative of a basic number instinct but they provide little insight about the nature of the underlying representations. And it seems nearly impossible to find indicators of more advanced understanding in trained animals that could not be explained by some simpler non-numerical process. There is a confusing kind of circularity in the use of some indicators, while others are highly vulnerable to over-interpretation. Complicating matters is the fact that researchers cannot agree on definitions for even the most basic numerical concepts.

In summary, the two questions posed at the beginning of this paper have not yet been satisfactorily answered. It is evident that non-human primates are sensitive to numerosities and can respond to this property of sets. It is also clear they can learn to map symbols onto stimuli on the basis of numerosity. If this were not the case it would be impossible to train them to label and order numerical arrays, which a variety of species have plainly learned to do. What is still unknown is the extent of their understanding about what numbers mean. It is possible they possess some degree of insight, but the indicators so far have not allowed us to define what the limits of that insight might be.

References

- Biro, D., & Matsuzawa, T. (2001). Chimpanzee numerical competence: cardinal and ordinal skills. In Matsuzawa, T. (Ed.), *Primate origins of human cognition and behavior* (pp. 199-225). NY: Springer-Verlag.
- Boysen, S. T., & Berntson, G. G. (1989). Numerical competence in a chimpanzee (*Pan troglodytes*). *Journal of Comparative Psychology*, *103*(1), 23-31.
- Boysen, S. T., & Berntson, G. G. (1990). The development of numerical skills in the chimpanzee (*Pan troglodytes*). In S. Taylor, K. R. Gibson, et. al. (Eds.), *"Language" and intelligence in monkeys and apes: comparative developmental perspectives* (pp. 435-450). NY: Cambridge University Press.
- Boysen, S. T., & Berntson, G. G. (1995). Responses to quantity: perceptual vs. cognitive mechanisms in chimpanzees (*Pan troglodytes*). *Journal of Experimental Psychology: Animal Behavior Processes*, *21*, 83-86.
- Boysen, S. T., Berntson, G. G., Hannan, M. B., & Cacioppo, J. T. (1996). Quantity-based interference and symbolic representations in chimpanzees (*Pan troglodytes*). *Journal of Experimental Psychology: Animal Behavior Processes*, *22*(1), 76-86.
- Boysen, S. T., Berntson, G. G., Shreyer, T. A., & Hannan, M. B. (1995). Indicating acts during counting by a chimpanzee (*Pan troglodytes*). *Journal of Comparative Psychology*, *109*, 47-51.
- Boysen, S. T., & Hallberg, K. I. (2000). Primate numerical competence: contributions toward understanding nonhuman cognition. *Cognitive Science*, *24*(3), 423-443.
- Brannon, E. M., & Terrace, H. S. (1998). Ordering of the numerosities 1 to 9 by monkeys. *Science*, *282*, 746-749.
- Davis, H., & Perusse, R. (1988). Numerical competence in animals: definitional issues, current evidence, and a new research agenda. *Behavioral and Brain Sciences*, *11*, 561-615.
- Dehaene, S. (1997). *The number sense*. New York: Oxford University Press.
- Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition*, *44*, 43-74.
- Hauser, M. D. (1997). Tinkering with minds from the past. In *Characterizing human psychological adaptations, CIBA foundation symposium 208* (pp. 95-129). New York: John Wiley and Sons.
- Hauser, M. D., MacNeilage, P., & Ware, M. (1996). Numerical representations in primates. *Proceedings of the National Academy of the Sciences, USA*, *93*, 1514-1517.
- Kaufman, E. L., Lord, M. W., Reese, T. W., & Volkman, J. (1949). The discrimination of visual number. *American Journal of Psychology*, *62*, 498-525.
- Kawai, N. (2001). Ordering and planning in sequential responding to Arabic numerals by a chimpanzee. *Psychologia*, *44*, 66-69.
- Meck, W. H., & Church, R. M. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behavior Processes*, *9*(3), 320-334.
- Olthof, A., Iden, C. M., & Roberts, W. A. (1997). Judgments of ordinality and summation of number symbols by squirrel monkeys (*Saimiri sciureus*). *Journal of Experimental Psychology: Animal Behavior Processes*, *23*(3), 325-339.

- Pepperberg, I. M. (1987). Evidence for conceptual quantitative abilities in the african grey parrot: labeling of cardinal sets. *Ethology*, *75*, 37-61.
- Rumbaugh, D. M., Savage-Rumbaugh, E. S., & Hegel, M. T. (1987). Summation in the chimpanzee (*Pan troglodytes*). *Journal of Experimental Psychology: Animal Behavior Processes*, *13*, 107-115.
- Rumbaugh, D. M., & Washburn, D. A. (1993). Counting by chimpanzees and ordinality judgments by macaques in video-formatted tasks. In S. T. Boysen & E. J. Capaldi (Eds.). *The development of numerical competence: animal and human models* (pp. 87-108). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Sulkowski, G. M., & Hauser, M. D. (2001). Can rhesus monkeys spontaneously subtract? *Cognition*, *79*, 239-262.
- Suzuki, K., & Kobayashi, T. (2000). Numerical competence in rats (*Rattus norvegicus*): Davis and Bradford (1986) extended. *Journal of Comparative Psychology*, *114*(1), 73-85.
- Terrell, D. F., & Thomas, R. K. (1990). Number-related discrimination and summation by squirrel monkeys (*Saimari sciureus sciureus* and *S. boliviensis boliviensis*) on the basis of the number of sides of polygons. *Journal of Comparative Psychology*, *104*(3), 238-247.
- Thomas, R. K., & Chase, L. (1980). Relative numerosness judgments by squirrel monkeys. *Bulletin of the Psychonomic Society*, *16*, 79-82.
- Thomas, R. K., Phillips, J. A., & Young, C. D. (1999). Comparative cognition: human numerosness judgments. *American Journal of Psychology*, *112*(2), 215-233.
- Tomonaga, M., & Matsuzawa, T. (2000). Sequential responding to Arabic numerals with wild cards by the chimpanzee (*Pan troglodytes*). *Animal Cognition*, *3*, 1-11.
- Uller, C., & Hauser, M., & Carey, S. (2001). Spontaneous representation of number in cotton-top tamarins (*Saguinus oedipus*). *Journal of Comparative Psychology*, *155*(3), 248-257.
- Wynn, K. (1998). Numerical competence in infants. In Donlan, C. (Ed.), *The development of mathematical skills* (pp. 3-25). Psychology Press.