7-15-2008

Numerical Cognition in Rhesus Monkeys (Macaca mulatta)

Emily Harris Marr
Georgia State University

NUMERICAL COGNITION IN RHESUS MONKEYS (*MACACA MULATTA)*

by

Emily Harris Marr

Under the Direction of Dr. David A. Washburn

ABSTRACT

Over the past few decades, researchers have firmly established that a wide range of nonhuman animals exhibit some form of numerical competence. The focus of this research was to define further the extent of numerical ability in rhesus monkeys, and specifically to determine whether the animals possess a symbolic understanding of Arabic numerals. This required examining the stimulus attributes (e.g., number vs. hedonic value) represented by the numerals, as well as the precision (e.g., absolute vs. relative) and generality of those representations. In chapters 2 and 3, monkeys were required to compare and order numerals and were rewarded with either proportional or probabilistic rewards. The results indicated that monkeys were relying on the ordinal or absolute numerical values associated with each numeral and not hedonic value or learned 2-choice discriminations. The studies in chapters 4 and 5 indicated that monkeys can use numerals to symbolize an approximate number of sequential motor responses. The study in Chapter 6 tested the generality of the monkeys' symbolic number concept using transfer tests. The results indicated that some monkeys are able to abstract number across presentation mode, but this ability is only exhibited under limited conditions. Collectively, these studies provide evidence that rhesus monkeys view Arabic numerals as more than sign-stimuli associated with specific response-reward histories, but that numerals do not have the same precise symbolic meaning as they do for humans.

INDEX WORDS: Monkeys, *Macaca mulatta*, numbers, symbols, numerical ability

NUMERICAL COGNITION IN RHESUS MONKEYS (*MACACA MULATTA)*

by

EMILY HARRIS MARR

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

in the College of Arts and Sciences

Georgia State University

2008

NUMERICAL COGNITION IN RHESUS MONKEYS (*MACACA MULATTA)*

by

EMILY HARRIS MARR

Committee Chair: Dr. David Washburn

Committee: Dr. Michael Beran

Dr. Michael Owren

Dr. Rose Sevcik

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

Summer 2008

ACKNOWLEDGMENTS

I thank my advisor, Dr. David Washburn, for his guidance, collaboration, and encouragement over the years. I also thank my committee members, Dr. Michael Beran, Dr. Michael Owren, and Dr. Rose Sevcik, for their commitment to this project. I am also grateful to Michael Beran for writing the computer software that was used in the majority of these experiments. Lastly, I thank my parents, my husband Bo, and my wonderful graduate student friends for their support during this process.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1: Literature review

In his 1945 book on the role of mathematics in civilization, Eric Bell wrote, "Even stranger things have happened; and perhaps the strangest of all is the marvel that mathematics should be possible to a race akin to the apes" (p. 594). Whereas it is amazing that human beings, a species sharing common ancestry with the chimpanzee and gorilla, have developed such elegant and complicated systems of formal mathematics, it is arguably more interesting that apes themselves have shown impressive numerical reasoning abilities.

The topic of numerical competence in nonhuman animals (hereafter "animals") has fascinated researchers for over a century, but early demonstrations of numerical aptitude were often plagued by lack of scientific controls. The most infamously flawed display of animal numerical prowess involved a horse named Clever Hans who lived in Germany in the early 1900s and was rumored to be able to add, subtract, multiply, and divide Arabic numerals. When it was discovered that the horse was not solving the numerical problems, but instead simply responding to small, unintentional cues from its trainer, it cast a shadow of suspicion over the entire area of research (Davis & Memmott, 1982). The positive aspect of this controversial history is that subsequent investigators were more conscious of the need for rigorous experimental controls when conducting studies of animal numerical abilities and the need for caution when interpreting the results.

These subsequent studies have firmly established that a variety of animals ranging from rats (*Rattus norvegicus*; e.g., Burns, Goettl, & Burt, 1995; Church & Meck, 1984; Davis & Albert, 1986; Platt & Johnson, 1971) to dolphins (*Tursiops truncates*; e.g., Kilian, Yaman, von Fersen, & Güntürkün, 2003; Mitchell, Yao, Sherman, & O'Regan, 1985) to chimpanzees (*Pan troglodytes*; e.g., Beran, 2001, 2004; Boyson & Berntson, 1989, 1995; Matsuzawa, 1985, Murofushi, 1997; Rumbaugh, Savage-Rumbaugh, & Hegel, 1987) possess a sensitivity to number. It is possible that number is not a particularly salient cue for animals and will be used only as a "last resort" when alternative cues are not available (Breukelaar & Dalrymple-Alford, 1998; Davis & Bradford, 1986; Davis & Memmott, 1982), but there is considerable evidence that animals routinely and automatically respond to numerical value as readily as they respond to perceptual properties such as shape and color (Cantlon & Brannon, 2007; Capaldi & Miller, 1988; Dehaene, 1997; Gallistel, 1993).

Associative models of numerical cognition

Stimulus control by numerousness

Although the question of whether animals can respond to stimulus numerousness has been answered, it remains unclear whether this reflects conceptual knowledge about numerousness (a so-called "concept of number") or conversely is explicable via low-level perceptual and associative mechanisms. The radical behaviorist finds behavioral control by the numerousness attribute of stimuli to be no different than stimulus control by other stimulus dimensions (e.g., shape, color, size). Skinner (1963), for example, argued that all behavior can be explained as a function of environmental histories and reinforcing consequences. Rewarding an animal for a particular response increases the likelihood that the response will occur again in the future. According to this view, stimulus quantities, and symbols that represent quantities, function as fixed parts of a stimulus-response-reward association. In other words, the animal learns the specific response that must be executed in the presence of the numerical stimulus, or combination of numerical stimuli, in order to obtain a reward within the confines of that specific task. This perspective is illustrated in more recent discussions by Almeida, Arantes, and Machado (2007); Machado and Rodrigues (2007); McGonigle (1988); Mechner and Guevrekian

(1962); Silberberg and Fujita (1996); and Smirnova, Lazareva, and Zorina (2000). As Nevin (1988) asked rhetorically, "Is our science more effectively advanced by debating rules for the use of terms like 'counting' that imply active but unobservable organismic processing, or by identifying and quantifying variables that control behavior?" (p. 595).

Subitizing and other perceptual explanations

C. Lloyd Morgan (1914), an early critic of animal numerical research, espoused a similar view that animals were limited to instinctive behavior, sensory experience, and associative learning, rather than higher-order cognitive processes such as concept formation. He argued that nonhuman numerical behavior could be explained in terms of the "simpler" process of timing. This untested assumption that timing reflects a more basic mechanism for explaining number-related judgments, was echoed in the influential literature review by Davis and Pérusse (1988).

Another explanation of human infant and nonhuman numerical ability which relies on relatively low-level perceptual variables is subitizing. The term subitizing was originally coined to refer to the rapid and near-parallel identification of small quantities of stimuli by adult humans. Studies show that when adult humans are asked to identify the number of items flashed on a screen, most will identify 1, 2, 3 or 4 stimuli with little difference in accuracy and latency. For arrays of about five or more stimuli, however, response time increases linearly with array size at a rate of about 200-300 ms per item (Jensen, Reese, & Reese, 1950; Klahr, 1973; Mandler & Shebo, 1982; Trick & Pylyshyn, 1994). In addition, when stimuli are presented under degraded conditions (e.g., flashed rapidly on a screen or masked after presentation), accuracy begins to decline at approximately 5 items (Atkinson, Campbell, & Francis, 1976; Oyama, Kikuchi, & Ichihara, 1981; Simon & Vaishnavi, 1996).

Although the discontinuous response function has been replicated in a variety of studies, researchers are still unclear as to the exact nature of the subitizing process. Many researchers believe that subitizing reflects a perceptual variable, such as pattern-matching or item-grouping, and does not reflect counting, enumeration, estimation, or other aspects of cognition that adults use to quantify larger displays of items (Gelman & Gallistel, 1978; Logan & Zbrodoff, 2003; Wender & Rothkegel, 2000; Wolters, van Kempen, & Wijlhuizen, 1987). For example, Mandler and Shebo (1982) argued that the fast reaction times for arrays of 1 to 3 items are based on learned canonical patterns. One dot forms a point, two dots can be connected by a line, and three dots typically fall in a triangular configuration. The display acts as a retrieval cue which activates a numerical response that has been associated with a similar arrangement of items through experience.

In contrast, Trick and Pylyshyn assert that the process of subitizing exploits a limited-capacity, preattentive mechanism for individuating a small number of items (Pylyshyn, 1989, 2001; Trick & Pylyshyn, 1994). This mechanism, which they refer to as the FINST mechanism, operates after the spatially parallel processes of feature detection and grouping, but before the serial processes of spatial attention. Trick and Pylyshyn argue that FINSTs, like human pointing fingers, allow one to select certain items for attentional processing without providing information about the item properties.

In another theory, Harris and Washburn (2008) proposed that subitizing is related to the cognitive concept of visual sensory storage, which is often called the “iconic buffer” or “icon.” The icon provides persistent access to visual stimuli and can be scanned by attention according to physical cues (location, color, size, and so forth). According to this theory, approximately four items can be identified quickly and comparably because that is the maximum number that can be scanned with attention before the icon has faded.

Subitizing and other direct perceptual processes were originally used by some researchers to explain all nonhuman animal and human child numerical abilities. This tendency was based on the observation that performance for both of these groups appeared to break down at the point where adults shift from subitizing to counting (approximately four items; Davis & Pérusse, 1988; Gast, 1957). We now know, however, that animals and infants are capable of performing numerical tasks with stimuli that far exceed the classic subitizing range (e.g., Beran, 2001, 2004; Brannon & Terrace, 1998; 2000; Cantlon & Brannon, 2006; Lipton & Spelke, 2003; Matsuzawa, 1985; Mechner, 1958; Olthof, Iden, & Roberts, 1997; Washburn & Rumbaugh, 1991; Xu & Spelke, 2000). This evidence means that subitizing can only be used to explain a portion of infant and animal numerical behavior. For example, Murofushi (1997) trained a chimpanzee to label sets of objects with Arabic numerals as large as seven. Results revealed that the chimpanzee's reaction time function was nearly flat for one to three objects and then generally increased as the number of objects increased. Based on this pattern, Murofushi concluded that the chimpanzee subitized when presented with one to three objects and relied on another process known as analogue magnitude estimation when presented with higher numbers.

Although subitizing is usually described as a visual process, Davis and Pérusse (1988) argued that it may also form the basis of sequential numerical discriminations. According to this hypothesis, sequential subitizing, also called rhythm discrimination, is based on recognizable rhythmic patterns, just as simultaneous subitizing is based on visual patterns. Thus, subitizing is a possible mechanism for both simultaneous and sequential discriminations involving a small number of items.

Representational models of numerical cognition

In contrast to the behaviorist view, cognitive interpretations of nonhuman numerical ability propose that animals have concepts of number (e.g., of “threeness”, “fourness”, and so forth) that emerge from experience and guide behavior. Conceptual behavior has been defined in the literature as correct responses “which do not depend upon prior experience with the specific stimuli being presented" (Thomas, 1984, p. 650). Thomas (1988) distinguished studies of numerical competence that examine conceptual processes from studies in which low-level rote memorization of particular stimulus exemplars can be used for responding. According to the conceptual view, animals respond correctly in numerical tasks not just because they have been conditioned to respond in a certain way when presented with specific numerical stimuli, but because they have a concept of number that can be used in (i.e., transferred to) a variety of tasks. Thomas and Lorden (1993) cautioned that although there are enough carefully controlled studies to conclude that some animals are capable of conceptual use of number, a “conservative and cautious view of animals’ use of number is warranted” because many studies do not properly control nonnumerical cues that can be used as the basis for responding (p. 129).

Analog magnitude model

Other researchers, such as Gelman and Gallistel (1978), are strong champions of the conceptual view of animal numerical ability and argued forcefully against the tendency to underestimate the cognitive abilities of preverbal children and nonverbal animals. In response to Morgan’s chastisement of researchers who invoke higher-order mechanisms when simpler explanations of numerical behaviors are possible, Davis and Memmott (1982) issued a contrasting warning that researchers should not “exclude on a priori grounds the possibility of control by seemingly more complex stimulus dimensions” (p. 549). The analog magnitude model of nonverbal numerical ability provides support for the argument that number is different from other discriminatory aspects of stimuli. According to this model, animals form cognitive representations of number using a mental accumulator mechanism instead of relying on direct perceptual processes. Behavior is then based on this numerical cognition, reflecting the numerical categories (or the symbols that represent the mental categories) that are mapped onto this accumulator (Gallistel & Gelman, 2000; Meck & Church, 1983).

The accumulator mechanism was originally proposed by Gibbon (1981) to explain timing, but was later modified by Meck and Church (1983) and Gallistel and Gelman (2000) to account for nonverbal enumeration. According to this model, organisms that display numerical ability have an internal pacemaker in the brain that emits a stream of pulses at a steady rate and a switch gate that can be opened and closed to allow pulses to enter an accumulator. The pacemaker can be used as a timing device if the gate is left open during an event or it can be used as a counter if the gate is opened and closed whenever an object or event is enumerated. At the end of the timing or counting process, the accumulator sums the impulses that were allowed to enter and transfers that magnitude to working memory. The animal makes a judgment by comparing the current value in the accumulator to values stored in long-term memory. According to this model, animals trained with Arabic numerals, or other numerical symbols, learn decision rules that allow them to assign a specific symbol to a range of accumulator values (Dehaene, 1997; Gallistel & Gelman, 1992).

An important aspect of the accumulator model is that the memory for magnitudes is imperfect and defined by scalar variability. This means that trial-to-trial variability in responding increases in direct proportion to the quantity represented. Dehaene (1997) likened the accumulator to a beaker that gradually fills with water as items are counted and described the variability as sloshing in the beaker. The sloshing introduces noise into the recalled magnitudes, with increasing amounts of noise for larger numbers. In other words, the value obtained from reading the same memory repeatedly varies from reading to reading, with increasing variability for larger numbers (Gallistel, 1999).

Evidence for these noisy representations can be seen in many numerical studies involving animals. For example, when rats (Mechner, 1958; Platt & Johnson, 1971), pigeons (*Columba livia*; Xia, Emmerton, Siemann, & Delius, 2001; Xia, Siemann, & Delius, 2000), and chimpanzees (Beran and Rumbaugh, 2001) are required to make a certain number of responses, there is a proportional increase in the variability of responding with increases in target number. Additional evidence for the approximate and variable nature of numerical representations can be seen in studies requiring relative numerousness judgments. If the subjects are using analog magnitude representations then judgments of numerical inequality, like judgments of physical magnitudes such as weight and length, should obey Weber's law. This law states that the discriminability of two perceived magnitudes is determined by the ratio of their objective magnitudes. This is because the degree of overlap between representations remains constant when the ratio of the means is held constant. The larger the ratio of two subjective magnitudes, the greater the overlap in the signal distributions, hence the more difficult it is to discriminate the signals from the two different numbers (Gallistel & Gelman, 1992; Nieder & Miller, 2004; Whalen, Gallistel, & Gelman, 1999).

Studies assessing the numerical abilities of nonhuman primates (e.g., Beran, 2001, 2004; Flombaum, Junge, & Hauser, 2005; Jordan & Brannon, 2006; Lewis, Jaffe, & Brannon, 2005; Rumbaugh et al., 1987) and human infants (e.g., Lipton & Spelke, 2003; Xu & Spelke, 2000; Xu, Spelke, & Goddard, 2005) have found evidence of a ratio limit as predicted by Weber's law.

Behavior that conforms to Weber's law also exhibits distance and magnitude effects. This means that responding increases in accuracy as the numerical distance between comparison items increases and responding decreases in accuracy and becomes more variable as the magnitude of the quantities increase. These effects have been reported in many studies of animal numerical ability (e.g., Anderson, Stoinski, Bloomsmith, Marr, Smith, & Maple, 2005; Boysen & Berntson, 1995; Brannon & Terrace, 1998, 2000; Judge, Evans, & Vyas, 2005; Smith, Piel, & Candland, 2003).

In recent years, researchers have found a possible neural basis for the observed distance and magnitude effects. Nieder, Freedman, and Miller (2002) and Sawamura, Shima, and Tanji (2002) recorded number-sensitive neurons in the parietal and prefrontal cortices of rhesus monkeys (*Macaca mulatta*) and found that each cell showed peak activity for one quantity and a systematic reduction of activity as the presented quantity varied from the optimal quantity. These overlapping tuning curves could produce the distance effect observed in humans and animals. The neurons also became less precisely tuned as their preferred quantity increased, which could be the basis of the magnitude effect.

There are two major cognitive models proposed to explain distance and magnitude effects and they differ in the form that the quantity representation takes. The widely known logarithmic-compression hypothesis assumes that the subjective number continuum is logarithmically compressed so that the representations of 10 and 11 lie closer together on a mental number line than the representations of 2 and 3 (Dehaene, 1997, 2003; Dehaene & Changeux, 1993; Dehaene & Mehler, 1992; Moyer & Landauer, 1967). According to this theory, it is more difficult to discriminate higher numbers because they are subjectively closer together. An alternative hypothesis is that the subjective number continuum is linear and positions farther along the continuum are less precisely located (Brannon, Wusthoff, Gallistel, & Gibbon, 2001; Fetterman, 1993; Gallistel & Gelman, 1992; Gibbon, 1977). Thus, it is more difficult to discriminate higher numbers because their locations on the number line are more variable.

Regardless of whether the distance and magnitude effects are explained by a linear or logarithmic mental number line, the accumulator model makes the interesting prediction that there is no absolute limit to the numbers that can be discriminated. Research has shown that animals such as rhesus monkeys can discriminate sets as large as 20 and 30, given that the ratio is sufficiently small (Cantlon & Brannon, 2006). Human infants can also discriminate between large sets of visual objects and sounds, provided that the ratio of the two choices is small (Lipton & Spelke, 2003; Xu et al., 2005). For example, Xu and Spelke (2000) found that 6-month-old infants could discriminate between 8 and 16 elements (1:2 ratio), but not 8 and 12 elements (2:3 ratio).

The similarities between the numerical capacities of preverbal human infants and nonverbal animals have caused some psychologists to hypothesize that human mathematical ability shares an evolutionary past with the numerical abilities observed in animals (Butterworth, 1999; Dehaene, 1997; Gallistel & Gelman, 1992, 2000). Although humans demonstrate uniquely sophisticated numerical capacities such as division and calculus, the more basic numerical representations available to animals such as birds and monkeys may be the building blocks from which our mathematical ability was constructed over the course of evolution. Dehaene, Dehaene-Lambertz, and Cohen (1998) argued that number processing in humans and animals is based on a shared neural system that provides the foundation for higher-level arithmetic in humans.

Adult humans obviously have an advantage over human infants and nonhuman animals in terms of numerical ability because they can use number words and symbols to communicate precise quantities and perform formal operations such as multiplication and division. Evidence suggests, however, that knowledge of symbolic number systems does not eliminate the continuous quantitative mode of representation that we share with other animals. When presented with Arabic numerals, human brains automatically access the corresponding analog quantity representation, even when it is not necessary and may actually interfere with the task. As Dehaene (1997) wrote, “Our brain, like that of the rat, has been endowed since time immemorial with an intuitive representation of quantities” (p. 7).

Moyer and Landauer (1967) first discovered the automatic activation of a continuous representation by presenting adults with pairs of Arabic numerals and measuring the amount of time required for them to choose the largest digit. They found that responding was slower and less accurate for numerically close numbers such as 5 and 6, compared to more distant numbers such as 2 and 9. For equal distances, reaction time increased and accuracy decreased as the numbers became larger. The same thing happens when humans are asked to judge two physical magnitudes, such as weights or line segments.

The effects of continuous representations are also found for comparisons of two-digit numbers that can be performed simply by examining the first digit. For instance, when people are presented with the numbers 72 and 64, the simplest way to determine which number is larger is to compare the 7 and 6. However, experimental data show that the second digit affects reaction time. It takes more time to determine that 71 is larger than 64 compared to the time it takes to determine that 79 is larger than 64 (Dehaene, 1997). Distance effects are also seen in experiments requiring participants to judge whether digits are the same or different (Dehaene &

Akhavein, 1995). The presence of distance effects, even when numbers are presented in symbolic form, suggests that the human brain converts numbers internally from the symbolic format to a continuous, quantity-based, analogical format. This immediate and unconscious conversion is beneficial when we are asked to estimate the outcome of an operation such as 64 + 132, but creates accuracy and latency costs in simple comparison tasks.

A phenomenon known as the SNARC effect also suggests that Arabic numerals automatically evoke a non-symbolic representation (Dehaene, Bossini, & Giraux, 1993; Dehaene & Mehler, 1992). When subjects are shown the digits 0 through 9 and asked to respond with a left-hand or right-hand key depending on whether the digit is odd or even, larger numbers produce faster responses with the right hand and smaller numbers produce faster responses with the left hand. This result suggests that participants possess an internal number line with a left-to-right orientation.

Humans can also rely on this non-symbolic system of number processing for estimating and combining sets of elements when exact representations are not necessary or possible. In a study by Barth, La Mont, Lipton, Dehaene, Kanwisher, and Spelke (2006), adults performed numerical comparison and addition tasks involving large arrays of dots. The arrays were flashed quickly on a screen to prevent the use of formal counting. Results revealed that accuracy depended on the ratios of the compared quantities, which is a hallmark of approximate representation. An additional experiment revealed that 5-year-old children with no relevant knowledge of symbolic arithmetic could also perform the visual addition and comparison tasks at above-chance levels.

Whalen et al. (1999) adapted the procedure that Mechner (1958) and Platt and Johnson (1971) used with rats to study nonverbal counting in adult humans. In one task, participants attempted to produce a target number of key presses at a rate that made vocal or subvocal counting unlikely. In another task, participants estimated the number of flashes in a rapid sequence. In both tasks the participants were instructed not to verbally count, but to approximate the correct answer. The mean estimates in both tasks were proportional to the target values and variability increased as the targets increased. This is the same pattern of scalar variability obtained with animals in similar studies and supports the idea that adults and other animals share a nonverbal magnitude system for representing number.

In a study by Beran, Taglialatela, Flemming, James, and Washburn (2006), adults were repeatedly shown two sets of sequentially presented items of varying sizes and were asked to repeat the alphabet out loud during the trials to prevent them from counting. Results revealed that participants were able to choose the larger quantity at levels greater than chance performance, but accuracy decreased and variability in responding increased with increasing set size. These results provide further evidence that adult humans use approximate representations of numerosity when precise representations are not possible.

Based on results of these and other numerical studies with humans, Dehaene (1992) proposed a triple-code model of human numerical cognition. According to this model, adults represent numbers in an Arabic, verbal, or analogue magnitude code depending on the task. Neural evidence showing a double dissociation between the processing of Arabic and verbal numerals and the processing of quantity representations supports this model (Dehaene et al., 1998; McCloskey & Caramazza, 1987; Warrington, 1982).

Object-file model

In addition to the evidence that infants, animals, and adults have an approximate, ratio-dependent number system, there also is evidence for a more precise and limited nonverbal number mechanism. The object-file model explains this evidence by postulating that infants and animals form a representation of each individual object in the set when performing numerical tasks. According to the model, a separate object-file is opened for each object encountered so there is a one-to-one correspondence between objects and files. This means that representations are discrete and not approximate. The object-file model predicts that representations are limited to 3 or 4 items because that is the maximum number that can be simultaneously tracked by visual attention (Uller, Carey, Huntley-Fenner, & Klatt, 1999). The proposed reliance of both subitizing and the object-file model on object individuation has led some researchers to argue that the object-file model may be a specific case of subitizing (Brannon & Roitman, 2003). The prediction of an absolute numerical limit contrasts with the analog magnitude model, which predicts that performance is limited by numerical ratio and not the absolute number of items.

The majority of support for the object-file model is found in the infant literature. Infants that can easily discriminate 1 versus 2 items and 2 versus 3 items, fail when tested with comparisons of 3 versus 4 and 3 versus 5 (Antell & Keating, 1983; Starkey & Cooper, 1980; Strauss & Curtis, 1981; van Loosbroek & Smitsman, 1990). In a more recent study involving a manual search task, Carey (2004) found that infants succeeded at representing 1, 2, and 3 hidden objects, but failed to differentiate 4 from 2. Feigenson, Carey, and Hauser (2002) found that 10- and 12-month-old infants who watched experimenters place crackers into two opaque boxes succeeded in choosing the box with the larger number of crackers for comparisons of 1 versus 2 and 2 versus 3, but showed no preference for comparisons of 2 versus 4, 3 versus 4, and 3 versus 6. It must be noted, however, that the infants could have been using the total amount of food rather than the number of crackers to make their choice. In fact, when the amount of food was equated (a choice between 1 large and 2 small crackers), the infants showed no preference. If the infants in these studies were discriminating based on numerousness, the data suggest an upper limit of approximately 4 objects, which contrasts with infant studies showing a ratio limit rather than an absolute set size limit (e.g., Lipton & Spelke, 2003; Xu et al., 2005; Xu & Spelke, 2000).

In addition to infant studies, there are several animal studies that provide support for the object-file model. In a study by Hauser, Carey, and Hauser (2000), monkeys watched experimenters place apples, one at a time, into opaque boxes. Controls were used so that the amount of time spent filling the containers was not correlated with the number of slices. The animals reliably selected the box with the larger number of apples for discriminations of 5 versus 4, but failed with discriminations of 4 versus 8 and 3 versus 8. These findings provide support for an absolute number limit rather than a ratio limit, but it is possible that the limit was due to attentional factors rather than the mental mechanism involved. The untrained monkeys performed the task in their natural habitat while surrounded by other group members. This distraction may have caused a decrease in accuracy for sequential presentations requiring sustained attention.

In a study by Pepperberg and Gordon (2005), a parrot (*Psittacus erithacus*) responded to sets of items with verbal number labels. Its errors were randomly distributed across all set sizes tested (0-5) and it made the same number of errors when presented with a set of five items as a single item. This behavior is not described well by the accumulator model, which predicts decreasing accuracy with increasing quantities. It is possible that tests with greater quantities will show a clear set-size limit. These studies contrast with the many animal studies showing ratio-dependent performance (e.g., Beran, 2001, 2004; Flombaum et al., 2005; Jordan & Brannon, 2006; Lewis et al., 2005; Rumbaugh et al., 1987).

These disparate results have prompted some researchers to suggest that animals and infants have two separate systems for representing number (Feigenson, Dehaene, & Spelke, 2004; Sulkowski & Hauser, 2001; Uller et al., 1999; Xu, 2003). An analog magnitude system may be evoked for representing a large number of items in an approximate manner and an object-file system may be evoked for representing a small number of items precisely. Furthermore, it is possible that certain methodologies prevent infants and animals from engaging the system for representing large numbers, which results in a set-size limit. Fias and Verguts (2004) argued that both systems have an evolutionary basis and together provide a more complete picture of nonverbal numerical ability.

Counting

The majority of researchers use terms such as “numerical ability,” “numerical competence,” or “numerical sensitivity” to describe the performance of animals in studies that involve some knowledge of number. The use of these broad terms leaves open the question of whether or not animals can count in the same sense that humans can count. Most researchers define true counting as a formal enumerative process that conforms to the principles proposed by Gelman and Gallistel (1978). True counting, unlike associative mechanisms proposed to account for numerical discrimination, involves mental representations of numerousness.

The first principle proposed by Gelman and Gallistel (1978) is “one-to-one correspondence,” which means that each item in the array corresponds to one and only one distinct tag. Adult humans use conventional number words as tags, but Gelman and Gallistel argued that it may be possible for nonlinguistic animals and prelinguistic human infants to count using nonverbal number tags. These unconventional tags, called “numerons,” may take any form as long as they are unique symbols that bear no physical relation to the items being tagged.

The second principle is "stable order," which means that the tags corresponding to each item in an array must be used in a consistent, reproducible sequence. In other words, a child using the idiosyncratic count sequence, "one, six, three," is truly counting so long as he or she uses the same sequence each time. The third principle is "cardinality," which means that the last tag used during the count sequence represents the value of the entire set. Children tend to make more mistakes when counting than adults do, but research shows that children younger than 3 years old can tag items in a stable, one-to-one manner and indicate the cardinal number of the set by repeating the last tag in the list (Gelman & Gallistel, 1978).

The last two principles, which are the "abstraction" and "order irrelevance" principles, concern the applicability of the other three counting principles. The abstraction principle means that any collection of items can be counted. Adult humans can count heterogeneous visible sets, auditory sequences, and even nonphysical constructs. Klahr and Wallace (1973) argued that children first learn to apply the counting procedure to objects that have similar perceptual features, but there is evidence that even young children can count heterogeneous sets of items (Fuson, Pergament, & Lyons, 1985; Gelman & Tucker, 1975) as well as actions and sounds (Wynn, 1990). The order irrelevance principle means that the order in which items in an array are counted is irrelevant. In other words, any of the count words can be assigned to any of the items in an array as long as the one-to-one correspondence principle is observed.

Although most researchers agree on this definition of true counting, there is still much controversy over where animal numerical ability fits into this structure. Some argue that the term "counting" should be reserved for humans (Davis & Pérusse, 1988), but several studies have provided evidence of animal behavior that conforms to one or more of the formal counting principles as defined by Gelman and Gallistel (1978). For example, research has shown that chimpanzees, like young children, tend to touch or point to each item when judging the number of items in an array (Boysen & Berntson, 1989; Boysen, Berntson, Shreyer, & Hannan, 1995; Boysen & Hallberg, 2000). These gestures, known as indicating acts, may help the child or animal coordinate the tagging process involved in the application of the one-to-one correspondence principle. Although evidence of indicating acts is not necessary to conclude that an animal is adhering to counting principles, it does suggest that the animal is using an enumerative process indicative of formal counting.

In other studies, animals trained to associate Arabic numerals with the corresponding number of food pellets were correctly able to order arrays of up to five numerals (Beran et al., 2008; Washburn & Rumbaugh, 1991). This behavior suggests that they understood the order of the symbols, as required by the stable order principle. In addition, there is evidence that chimpanzees can reliably apply the correct Arabic numeral to arrays of familiar and novel objects. This evidence suggests that they understand the special status of the last number in a count sequence, as described by the cardinal principle (Boysen & Berntson, 1989; Matsuzawa, 1985; Murofushi, 1997). Chimpanzees can also contact a number of dots on a computer screen equal to an Arabic numeral cue and then indicate the end of the count by contacting the numeral. This behavior further demonstrates use of the cardinality principle (Beran, Rumbaugh, & Savage-Rumbaugh, 1998; Rumbaugh, Hopkins, Washburn, & Savage-Rumbaugh, 1989; Rumbaugh & Washburn, 1993). Furthermore, the chimpanzees in these studies did not always contact dots in the same order for the same target numerals, which indicates an understanding of the order-irrelevance principle.

Gelman and Gallistel (1978) originally envisioned the formal counting process as something quite different from the process performed by the accumulator mechanism proposed by Meck and Church (1983). According to the original theory, formal counting required mental tags in the form of discrete symbols rather than analog magnitudes. In recent years, however, researchers including Gallistel and Gelman (1992) and Dehaene (1997) have argued that the operation of the accumulator mechanism conforms to the first three counting principles and thus may provide the basis for true counting. According to this recent theory, the accumulator conforms to the one-to-one correspondence principle because the gate allows a burst of pulses to be emitted into the accumulator once and only once for every item in the to-be-enumerated set. The order of different accumulator states is stable from one count to the next because the magnitude of the accumulator is proportional to the number of items in the set. Also, it conforms to the cardinal principle because the final state of the accumulator is used to represent the value of the whole set.

Despite attempts to fit animal behavior ascribed to an accumulator mechanism into the definition of formal counting, it is obvious that animals do not count in the same way as adult humans do. When adults count, they use a precise sequence of number words or symbols that result in an exact value. This behavior indicates a more precise representation than the fuzzy representations produced by an accumulator relying on inexact analog magnitudes. In fact, scalar variability, which is common in animal numerical studies, is typically not seen in older children who can count beyond 10 (Le Corre & Carey, 2007). Based on these differences, animal researchers are typically cautious about applying the term “counting” to animals that may be relying on representational mental processes very different from those humans use.

Determining the nature of numerical cognition

The associative and representational theoretical orientations are contrasted in the subsequent sections and chapters of this dissertation. Numerical stimuli, and particularly Arabic numerals that represent quantities, were presented using various paradigms designed to determine whether rhesus monkeys, like humans, truly understand the concept of number and the symbolic nature of Arabic numerals. To establish that the monkeys have conceptual knowledge of number it is necessary to confirm that behavior is based on the numerousness aspect of stimuli rather than perceptual variables or conditioned responses to specific stimuli. Some researchers argue that abstract numerical ability, established through cross-modal and cross-procedural transfer testing, is also essential to demonstrating a true concept of number (Davis & Pérusse, 1988; Seibt, 1982). In contrast to this view, Thomas and Lorden (1993) argued that evidence of generalization to trial-unique numerosities is sufficient evidence to conclude that an animal has a true concept of number, given that generalization is not based on physical properties of the stimuli and that all other perceptual and associative processes have been precluded. Regardless of whether generalization across procedure is included as one of the criteria for possessing a number concept, investigating this ability is important to defining the representations used by animals in numerical tasks. Another important aspect to defining animal numerical ability is determining the precision of those representations. Thus, discovering the nature of the monkeys' number concept required multiple experimental steps, including:

1. Determining the existence of a number concept by testing whether performance in numerical tasks is based on a representation of number or a nonnumerical variable, such as hedonic value or a learned matrix of 2-choice discriminations.
2. Determining the precision of the numerical representation by testing for ordinal versus absolute (cardinal) numerical knowledge. In other words, investigating whether monkeys perform numerical tasks using knowledge of the ordinal position of each numeral, or

whether monkeys, like humans, understand that numerals represent precise quantities of items or events.

3. Determining the generality of the monkeys' representations by testing their ability to abstract number across different presentation modes.

Taken together, these studies help to determine whether number is a stimulus attribute that comes to control behavior in a stimulus-response-reward associative fashion, or whether animals acquire a concept of number that then serves to guide behavior in various numerousness-relevant contexts. In addition, the studies provide information regarding the nature of the monkeys' numerical representations and their understanding of the symbolic nature of Arabic numerals.

To understand the different steps required to investigate the monkeys' concept of number, it is necessary to understand the benefits and limitations of the different paradigms commonly used in animal numerical research, as well as significant findings thus far. The next section of this chapter provides a comprehensive literature review of animal cognition research and describes how the studies presented in this dissertation were designed to address unanswered questions in the field.

Quantitative judgments

In order to conclude that animals understand the conceptual nature of numerical stimuli, it is necessary to show that the nonhuman subjects are responding to the numerical dimension of the stimuli and not other stimulus attributes (e.g., density, surface area, and configuration) that can be used as a cue to responding. This is important because these perceptual dimensions typically covary with number, and thus would produce the same response as number. In fact, some testing paradigms that are considered a part of the numerical cognition literature are actually designed to assess general quantitative skills and thus do not control for nonnumerical factors. In these studies, which are sometimes called “go for more” judgments, animals are required to choose between two or more visible quantities. The quantities are often composed of uniform food items, which minimizes the need for training because animals spontaneously choose the larger quantity. The use of uniform food items, however, means that these judgments could be based on density, volume, and surface area. Laboratory studies have shown that many species are able to recognize small differences between food quantities, but these judgments are most successful when the arrays are relatively small and the difference between the arrays is large. Animals in the wild are also adept at choosing the largest food quantity, which is not surprising given the obvious foraging benefits (Menzel, 1960).

In an example of a laboratory study involving quantitative judgments, Uller, Jaeger, Guidry, and Martin (2003) presented red-backed salamanders (*Plethodon cinereus*) with two clear containers housing differing numbers of fruit flies. The salamanders chose the container with the greatest amount of flies for discriminations of 1 versus 2, 2 versus 3, 3 versus 4, and 4 versus 6. Although the salamanders showed an ability to “go for more”, the choice may not have been based on number. The container with the greatest number of fruit flies also had the greatest volume of flies and the highest probability of having at least one fly moving around during the choice process.

Beran, Evans, Leighty, Harris, and Rice (2008) presented capuchin monkeys (*Cebus apella*) with two sets of 1 to 6 uniform cereal pieces. The sets were originally covered with opaque containers and the experimenters revealed the sets sequentially prior to selection. Thus, the two sets were never visible at the same time and the monkeys had to make selections without immediate visual access to the sets during responding. The monkeys were able to retain quantitative information in memory and make accurate discriminations. In a second experiment, the monkeys were presented with sets that had both visible and nonvisible food items in them at the time of the response, thus requiring the monkeys to sum the total amount of food that was available. The monkeys again succeeded, which indicates that capuchin monkeys are highly sensitive to differences in quantity.

The great apes have also been tested with differing food quantities and their abilities exceeded that of capuchin monkeys and salamanders. Dooley and Gill (1977) presented a chimpanzee named Lana with two quantities of cereal and allowed her to eat the quantity she selected. Lana chose the larger quantity at levels greater than chance for comparisons as large a 9 versus 10, although performance was better for small quantities that differed by more than a few items. In a similar study, a group of lowland gorillas (*Gorilla gorilla*) chose the larger of two food quantities at levels comparable to the chimpanzee (Anderson et al., 2005).

In a study by Beran, Evans, and Harris (2008), chimpanzees were presented with one set of graham crackers in a vertical orientation (stacked) and one set of graham crackers in a horizontal orientation. The animals showed some bias for choosing the set with the individually largest single food item, but in general they selected the set with the largest total amount of food. The ability of the chimpanzees to compare sets with different orientations illustrates the flexibility and proficiency of their quantitative comparison skills.

Animals in the wild need to be able to compare two quantities of food in order to survive, but summing quantities of food located in different places may also provide adaptive benefits. There is evidence from the laboratory that animals can perform some basic summation tasks. In a study by Rumbaugh et al. (1987), chimpanzees were allowed to choose between two food trays. Each food tray consisted of two food wells containing 1 to 5 chocolates. To select the tray with the greatest total quantity of chocolates, the chimpanzees had to sum across the spatially separated food wells and compare the two summed values. The animals chose the tray with the greatest value on more than 90% of trials and accuracy was related to the ratio of the sums being compared. The researchers suggested that the chimpanzees were not performing true addition, which involves enumerating the items in both sets and combining exact values. Instead, summation could have been accomplished by perceptually fusing the chocolates in adjacent food wells to create two total quantities for comparison. An addendum to this study provided strong evidence that the animals were summing the quantities and not simply avoiding the tray with the smallest single amount or selecting the tray with the largest single amount (Rumbaugh, Savage-Rumbaugh, & Pate, 1988).

Relative numerousness judgments with analog quantities

Despite the ability of various species to compare and sum differing food quantities, true numerical ability can only be assessed by controlling for nonnumerical features of the stimulus array including size of the items, total surface area, brightness, and placement of the items within the array. Judgments between two quantities differing in number are often referred to as relative numerousness judgments (RNJs), emphasizing the fact that the animal does not need to know the absolute value of either array to choose the larger or smaller of the two quantities.

Thomas, Fowlkes, and Vickery (1980) conducted a study of relative numerousness judgments in two squirrel monkeys (*Saimiri sciureus*) using simultaneously visible stimuli. The stimuli consisted of cards displaying an array of circles ranging in number from 2 to 8 and the monkeys were reinforced for choosing the smaller numerosity. To control for area, brightness, and specific pattern cues the researchers used three different sizes of circles and a variety of patterns for each numerical value. The monkeys were first trained with one pair of stimuli and the numerosities were increased progressively as they learned each combination. One monkey discriminated pairs up to 7 versus 8 circles and the other performed even better by successfully discriminating 8 versus 9 circles.

In another relative numerousness study, Terrell and Thomas (1990) presented squirrel monkeys with different irregular polygons and rewarded them for choosing the stimulus with the fewest sides. Two monkeys met a stringent 90% accuracy criterion when comparing polygons with 5 and 7 sides, one met criterion when comparing polygons with 6 and 7 sides, and one met criterion when comparing polygons with 7 and 8 sides. In a second experiment, two polygons were presented on each card and the sides had to be summed to determine which card had the total fewer sides. Three monkeys performed better than chance on comparisons of 6 and 8 sides and one monkey met the high accuracy criterion for 6 versus 8 and 7 versus 8, indicating that they were able to sum the sides of the polygons.

In addition to these studies of relative numerousness judgments in primates, there is evidence that other animals can discriminate between stimuli based on their relative numerosity. For example, pigeons presented with video displays consisting of small squares of red and blue elements learned to peck one side of the screen when the blue elements were more numerous than the red elements and the other side of the screen when the red elements were more numerous. Control conditions indicated that this behavior was based on the relative numerousness of the different colors and not other factors such as the spacing and size of the elements (Honig & Matheson, 1995). Pigeons were also able to discriminate the relative numerousness of items that differed in form such as X's and O's and images of birds and flowers (Honig & Stewart, 1989).

Relative numerousness judgments with symbols

These studies illustrate the systematic control conditions that are necessary to eliminate nonnumerical cues when analog stimuli are used to represent number. Stocklin (1999) has suggested that even with stringent control conditions there may be dimensions such as complexity that vary with number. Perhaps a stimulus consisting of 5 squares looks more complex than a stimulus consisting of 3 squares because of the extra corners and edges, regardless of the size of the squares or the density of the stimulus arrays. An alternative is to use arbitrary symbols, such as Arabic numerals, to represent number. The use of symbols eliminates the need for many of the control conditions used with analog stimuli. However, the lack of inherent numerical value means that the animals must be trained to associate the symbols with different numbers.

Mitchell et al. (1985) trained dolphins to associate 6 unique objects with differing numbers of fish. On each trial, dolphins were allowed to choose from an array of 2 or more objects and each object was consistently rewarded with a certain number of fish ranging from 0 to 5. After approximately 2,000 trials the dolphins learned to choose the available object that represented the greatest amount of fish for most pairings. This indicates that the dolphins were able to learn the relations among the stimuli.

In a series of studies, Boysen and her colleagues found evidence that the use of symbols can actually improve performance in tasks involving reverse contingency reinforcement. In one study, chimpanzees were trained to select among two different amounts of candy ranging from 1 to 6 pieces. The array that was chosen by the chimpanzee was then given to another chimpanzee observing the experiment. The active chimpanzee received the remaining nonselected array. Thus, the optimal strategy for the active chimpanzee was to choose the smaller array. The chimpanzees consistently selected the larger candy array even though it resulted in a smaller reward, but performance greatly improved when Arabic numerals were substituted for the candy arrays. In addition, performance increased as the numerical distance between the Arabic numerals increased even though the opposite was true for the candy arrays. The use of symbols allowed the chimpanzees to inhibit their natural instinct to choose the larger food array, thereby optimizing performance (Boysen & Berntson, 1995). This was also true when the nonselected choice was simply removed instead of given to an observer animal (Boysen, Berntson, Hannan, & Cacioppo, 1996) and when the chimpanzees were presented with mixed symbol-candy choices (Boysen, Mukobi, & Berntson, 1999). These results suggest that symbols can have an adaptive function because they represent number without exhibiting the same distracting properties as food items. It is possible that the adaptive benefit is limited to chimpanzees, however, because a similar study with orangutans (*Pongo pygmaeus*) found that the use of numerals did not increase performance (Shumaker, Palkovich, Beck, Guagnano, & Morowitz, 2001).

Despite the facilitative effects that Arabic numerals can have on some numerical tasks, a large number of training trials are usually required before the animal learns to select the numerals in the correct sequence (e.g., Olthof & Roberts, 2000; Washburn & Rumbaugh, 1991). Beran, Beran, Harris, and Washburn (2005) devised a system that facilitated rapid learning of the relations among symbols representing different quantities of food. They used colored plastic eggs as the symbols because color is a highly salient property for most nonhuman species. The use of eggs also ensured close spatial contiguity between the stimuli and the food items they represented because the food could be placed directly inside of the eggs. Results revealed that two chimpanzees and a rhesus macaque rapidly learned the relations between five colored eggs when all eggs of a given color contained a specific number of identical food items (e.g., blue eggs always contained four food pellets). However, all animals failed in a summation task, in which a single container was compared with a set of 2 containers of a lesser individual quantity but a greater combined quantity. This finding indicates that the animals had difficulty evaluating the sets of colored containers on the basis of more than one dimension (color and quantity).

Although the use of symbols and other stringent controls of nonnumerical features in these judgment tasks allow researchers to conclude that animals are sensitive to number, there are important limitations to this research paradigm. One limitation is that it is difficult to determine whether the judgments are based on the absolute or relative number of stimuli. For example, it is possible to judge that a set of 5 elements is greater in number than a set of 4 elements without knowing the exact number of elements in either set. If animals have a true understanding of the symbolic nature of Arabic numerals, their representations should include both relative and absolute numerical knowledge.

To test for absolute numeral knowledge, Beran et al. (2005) presented the primates in their study with trials involving one colored egg and one visible set of food items. All three animals performed at high levels, indicating that the animals had learned the approximate quantity of food items in eggs of a given color. Other researchers have devised alternative paradigms to investigate knowledge of absolute number in animals, which will be discussed later in this chapter.

Ordinal numerical judgments with analog quantities

Another limitation of relative numerousness studies is that some do not make it clear whether the animals understand the ordinality of these stimuli or whether they are simply learning pair-wise comparisons. Ordinality refers to the ability to place items in a particular sequence based on some quantitative property, such as size or number. The monkeys in the

Thomas et al. (1980) and Terrell and Thomas (1990) studies were trained with one pair of numerosities after another and could have performed discriminations such as 2 versus 3 and 3 versus 4 by learning that the numerosity 3 is rewarded when it is paired with 4, but not when it is paired with 2. This is not the same as knowing the ordered rule: 2 is less than 3, which is less than 4.

To conclude that animals are representing the numerical values associated with numerals and not simply responding to numerals as conditioned stimuli, it is necessary to determine that animals are not responding to numeral pairs based on learned 2-choice discriminations (e.g., if the stimuli are 4 and 5, pick 5 to get the reward; if the stimuli are 4 and 3, pick 4; and so forth). One way to show that an animal's performance is based on knowledge of numerical order is to train them on a range of numerosities and then test them on novel values outside of that range. Emmerton, Lohmann, and Niemann (1997) trained pigeons to differentiate visual arrays according to the relative number of their elements. Initially, the pigeons were reinforced for pecks to one response key when the stimulus depicted "many" elements (6 or 7) and were reinforced for pecks to a different response key when the stimulus depicted "few" elements (1 or 2). In subsequent tests, the novel intervening numerosities 3, 4, and 5 were introduced and nonnumerical factors such as brightness, size, shape, area, and contour of the elements were systematically controlled across tests. The pigeons were reinforced on all trials involving novel numerosities, regardless of their response. The investigators found that variations in performance corresponded to the values of the new numerosities. As the number of elements in each stimulus increased, so did the percentage of "many" choices made by the pigeons. The subjects responded to these novel numerosities as if they belonged on a continuum between the stimuli representing few and many. This indicates that they did not view the numerosities as disconnected categories, but instead recognized the relationship between them.

In another study involving transfer to novel numerosities, Kilian et al. (2003) trained a bottlenose dolphin to discriminate between two simultaneously presented stimuli differing in numerosity. The dolphin was trained with the numerosities 2 and 5 and then tested with nonreinforced trials involving all combinations of the novel numerosities 1 through 6. The dolphin showed evidence of immediate transfer for all of the novel numerosity pairs excluding 4 versus 5, which indicates that dolphins are also capable of representing ordinal relations among some numerosities.

In a frequently cited study of ordinal knowledge in animals, Brannon and Terrace (1998) trained two rhesus monkeys to respond to exemplars of the numerosities 1 to 4 in ascending order. The configuration of the exemplars was varied randomly between trials and the exemplars were controlled for nonnumerical features such as size, density, and color. The monkeys were then tested on their ability to order pairs of the numerosities 1 through 9. No reinforcement was provided for pairs involving novel numerosities and new exemplars were used on each trial so that the monkeys could not memorize the individual exemplars. Both monkeys responded in ascending order on each type of numerical pair, including those involving novel numerosities. As was true for many of the studies already discussed, accuracy was highest for small numerosities and those separated by a large numerical distance. Subsequent studies using similar procedures demonstrated that brown capuchin monkeys (Judge et al., 2005), as well as a hamadryas baboon (*Papio hamadryas*) and a squirrel monkey (Smith et al., 2003) could represent and order the numerosities 1 through 9 and rhesus monkeys trained to order the values

1 through 9 could correctly order pairs of the novel numerosities 10, 15, 20, and 30 (Cantlon & Brannon, 2006).

Although the two monkeys trained by Brannon and Terrace (1998) to order numerosities in ascending order (1-2-3-4) were correctly able to order the novel numerosities 5 through 9, it is interesting to note that another monkey trained to respond in descending order (4-3-2-1) did not exceed a chance level of accuracy on pairs of novel numerosities. This puzzling result prompted a follow-up experiment in which the researchers trained one rhesus monkey to respond to exemplars in a 4-5-6 order and another to respond to exemplars in a 6-5-4 order. Results revealed that the monkey trained on a 4-5-6 sequence was highly accurate when presented with novel pairs derived from the values 7 through 9, but performed below chance levels when presented with pairs derived from the values 1 through 3. The opposite was true for the monkey trained on a 6-5-4 sequence (Brannon & Terrace, 2000). These results suggest that the monkeys may not use an ordinal rule in the same way that humans would in a similar task. Instead, they may respond to test pairs based on which value is closer to the first value on which they were trained. This strategy still requires ordinal knowledge because the monkeys recognize which novel numerosities are closest to their initial training value, but it leaves open the question of whether or not monkeys can represent all of the values 1 through 9 on a continuous mental line.

Ordinal numerical judgments with symbols

Ordinal studies can also be conducted using symbols such as Arabic numerals instead of analog stimuli, but a different strategy is needed because novel symbols have no inherent meaning for the animals and therefore cannot be used to test the spontaneous transfer of an ordinal rule from one range of values to another. Instead, researchers test the animal's ability to transfer an ordinal rule to novel combinations of trained symbols that cannot be ordered using transitive inference.

In a study by Washburn and Rumbaugh (1991), two rhesus monkeys used a joystick to choose between pairs of the Arabic numerals 0 through 9 on a computer screen. Each selection was rewarded with the corresponding number of pellets (i.e., one pellet for choosing the numeral 1 and two pellets for choosing the numeral 2). During training the monkeys learned to choose the larger of the two numerals to receive the larger reward. The investigators withheld seven pairs during training and presented them later as novel probe trials. One monkey responded to these novel pairs of symbols at levels significantly greater than chance within the first few presentations. These probe pairings were carefully chosen by investigators so that they could not be solved by logical transitivity. For instance, the monkeys learned during training that 8 resulted in a larger reward than 7 and that 8 resulted in a larger reward than 6, but this information alone could not help them solve the probe pairing of 7 and 6. The monkeys may have solved the probe pairings using knowledge of the exact difference between pairs of numerals. In other words, during training they may have learned not only that 8 was greater than 7 and 6, but that 8 was greater than 7 by one pellet and 8 was greater than 6 by two pellets. A more parsimonious explanation is that the monkeys learned the absolute value of pellets associated with each numeral. For instance, they learned that choosing the numeral 8 resulted in eight pellets and choosing the numeral 7 resulted in seven pellets.

It is clear from the probe trials that the monkeys acquired some form of knowledge of the relative values of the numerals. In other words, they recognized that 8 > 7 and 7 > 6 and so forth for different pairs of numerals. However, this is not the same as knowing that all of the numerals are ordered in magnitude along a continuum such that 8 > 7 > 6. To investigate this question, Washburn and Rumbaugh (1991) presented the two monkeys with random arrays of up to five numerals. The monkeys were able to choose the numerals in descending numerical order, which indicates that the monkeys organized the numerals in a sequence according to their values.

In a series of similar experiments, Olthof and colleagues (Olthof et al., 1997; Olthof & Roberts, 2000) presented squirrel monkeys and pigeons with choices between all possible pairs of the Arabic numerals 0, 1, 3, 5, 7, and 9. Each numeral was rewarded with the corresponding number of food pieces. Both species learned to choose the larger numeral when presented with pair-wise comparisons and immediately chose the larger numeral when presented with novel pairs that could not be solved by transitive inference. They also learned to choose the largest numeral from a set of four numerals. In subsequent experiments, the animals were presented with stimuli that consisted of 1 to 3 numerals. They chose the largest total sum when presented with pairs of stimuli containing 2 numerals versus 2 numerals, 1 numeral versus 2 numerals, and 3 numerals versus 3 numerals, although performance was better for smaller ratios of total quantities. These results could not be explained by selection of the stimulus with the largest single numeral or avoidance of the smallest single numeral. This indicates that the animals summed the symbols that were located in close spatial proximity and that they had a fairly accurate representation of the individual quantities associated with each symbol.

Despite the success of animals trained to associate numerals with differing numbers of food rewards in the studies by Washburn and Rumbaugh (1991) and Olthof and colleagues (Olthof et al., 1997; Olthof & Roberts, 2000), questions remained regarding the interpretation of those data. First, presenting different numbers of food rewards based on differing numerals allowed for the possibility that the animals were responding on the basis of hedonic value. In other words, choices with higher numerical values were also associated with a greater quantity of food so responses may have been based on a judgment of how satiating a stimulus was instead of the numerical value of the stimulus.

There were also questions regarding the final experiment by Washburn and Rumbaugh (1991) in which 3, 4, or 5 numerals were presented at once on the screen. The monkeys exceeded chance levels of performance on those trials, suggesting that the monkeys had established an ordinal sequence based on those numerals and their magnitude. However, this final experiment occurred during the point at which the monkeys were most highly experienced in the task, and no control condition was offered to illustrate just how quickly the monkeys may have learned to respond to arrays of up to five completely novel stimuli. These comparison data would illustrate whether performance with larger arrays of numerals was supported by an ordinal representation of the numeral stimuli acquired during the previous experiments (in other words, whether the animals had established a competency in ordering numerals on the basis of their magnitudes) or was simply the result of rapid learning of a new ordinal sequence.

In a study by Beran et al. (2008), further aspects of ordinal learning with numerical stimuli were investigated in capuchin monkeys and rhesus monkeys. Both species were presented with pairings of the Arabic numerals 0 through 9, but some animals were given differential rewards based on the numeral selected and some were rewarded with a single pellet for every correct response. Both species learned to select the larger of the two numerals, and rhesus monkeys that were differentially rewarded performed above chance levels when presented with novel probe pairings. This provides evidence that performance was not the result of hedonic value that accrued because of reward contingencies. In a second experiment, the monkeys were first presented was arrays of 5 numerals (0-9) and then arrays of 5 letters. Both species performed better with the numerals, which indicated that performance was not just the result of rapid learning. Instead, an ordinal sequence of all stimuli had been learned during the first experiment.

Absolute numerical judgments with symbols

Despite the impressive evidence that these studies provide for ordinal knowledge in nonhuman primates, it is difficult to determine whether the judgments are based on the absolute or relative numerosity of the stimuli. As stated earlier, it is possible to judge that five elements are more than four elements without knowing the absolute number of either set of elements, which is also known as the cardinal value of the set. Animals in the wild typically confront situations where relative knowledge is sufficient. For example, it is important for an animal to know whether its allies outweigh its foes before engaging in a conflict, but it is not necessary to know the exact number of friends or foes. This fact led numerical researcher Hank Davis (1993) to conclude that absolute numerosity is a “distinctly human invention” (p. 109).

Despite the lack of ecological necessity, there is some evidence that the monkeys in Washburn and Rumbaugh (1991) learned the quantitative values represented by Arabic numerals. In a study of Stroop-like effects, six numeral-trained rhesus monkeys from the same laboratory learned to select the larger of two arrays of 1 to 9 letters (e.g., to select five As rather than four Cs; Washburn, 1994). When the arrays of letters were replaced with arrays of numerals, it was more difficult for the monkeys to choose the array with the most stimuli when that array was composed of the smaller numeral (e.g., four 1s versus two 5s) than when it was composed of the larger numeral (e.g., four 5s versus two 3s). These Stroop-like interference and facilitation effects suggest that these monkeys processed the absolute quantitative meanings of the numerical symbols automatically because of their prior training with these numerals, despite the fact that these meanings were irrelevant to the task.

In Experiment 1 of **Chapter 2: What do Arabic numerals mean to macaques?**, we further investigated the use of absolute versus ordinal representations by presenting six numeral-trained rhesus monkeys with pairs of Arabic numerals and pairs of dot arrays ranging in value from 1 to 9. Five of the monkeys received proportional rewards for every selection (e.g., picking the numeral 4 or an array of four dots netted four pellets) and the sixth monkey was rewarded with a proportional number of pellets for numerals, but probabilistic rewards for dot selections (e.g., correctly picking the larger array always netted one pellet). The alternative reward contingency was designed to address the criticism that rewarding animals with a number of food items corresponding to the value of the stimulus confounds number and hedonic value (Brannon & Terrace, 2002).

The monkeys were then given novel probe pairs involving one Arabic numeral choice and one dot array choice. If the monkeys had originally learned a complex matrix of values using knowledge of the relative difference and degree of relative difference between pairs of numerals, then they should be incapable of comparing symbols with actual quantities. Conversely, if the monkeys acquired knowledge of the absolute quantity of pellets represented by each Arabic numeral, then they might be able to compare symbols with analog dot arrays.

Results revealed that all of the monkeys were able to choose the largest value for probe trials involving one numeral and one dot array, even on the first exposure to these trials. This finding allowed us to rule out a matrix of learned values and also hedonic value as the basis for responding. The data therefore suggest that the monkeys had acquired knowledge about the absolute quantity of things represented by each Arabic numeral and could, even on probe trials, compare accurately this represented quantity to a visible array of dots.

An alternative possibility, however, is that performance reflected integration of two learned sequences instead of comparisons of quantity. Research indicates that monkeys trained to order two lists of four arbitrary stimuli (e.g., $A_1B_1C_1D_1$ and $A_2B_2C_2D_2$) immediately respond correctly at a greater than chance level when presented with comparisons of two items from different lists (e.g., A_1-C_2 and B_1-D_2; D'Amato & Colombo, 1988; Terrace, Son, & Brannon, 2003). It is possible that the monkeys in the Chapter 2 study perceived the numerals as one arbitrary list of stimuli and the dot quantities as another arbitrary list and ordered pairs of numerals and dots using only knowledge of their ordinal position.

In **Chapter 3: Ordinal-list integration for symbolic, arbitrary, and analog stimuli by rhesus macaques (*Macaca mulatta*)** we assessed whether the animals in our laboratory were responding to numerals on the basis of their ordinal value or whether they had a representation of the absolute value associated with each numeral. If Arabic numerals have ordinal value for the monkeys based on their previous numerical experience, they should learn to produce a list of Arabic numerals faster than they would learn to produce a list of unfamiliar arbitrary stimuli. Conversely, if past experiences using Arabic numerals have led to representations of those numerals that are linked to specific quantities (cardinal values), but not ordinal values, then the monkeys should show no advantage when learning to produce an ordinal list of numerals. In addition, these quantity representations should not lead to facilitative effects during integration of the numeral list with lists of arbitrary stimuli that are only associated with ordinal information. In Experiment 1 of Chapter 3, monkeys learned to order serially a list of five numerals, a list of five colored squares, and a list of five arbitrary signs. In Experiment 2, the monkeys received nonrewarded pair-wise comparisons of items from different lists, testing the ability of the monkeys to use ordinal position information to integrate the lists.

In addition to our investigation of the use of ordinal versus cardinal numerical knowledge, we also investigated the integration of analog quantities into ordinal lists. In Experiment 3 of Chapter 3, the monkeys received nonrewarded pair-wise comparisons of analog quantities and items from the three learned lists. Both of the monkeys involved in this study had previous experience viewing and responding to a variety of analog stimuli, so we hypothesized that the monkeys may spontaneously be able to integrate those types of stimuli into ordinal lists by converting the quantity information inherent in the analog sets into ordinal information. Together, the experiments in Chapter 3 provide us with a greater understanding of the nature of the monkeys' concept of number and their understanding of the symbolic nature of Arabic numerals.

The studies in Chapters 2 and 3 tested for absolute numerical knowledge using pairwise comparison tasks, but other paradigms exist as well. To test for absolute numerical representations, several researchers have attempted to train animals to match numerical quantities with specific symbols. For example, Ferster (1964) trained two chimpanzees to match the quantities 1 through 7 with binary numbers ranging from 001 to 111. In the first phase of training, the chimpanzees were shown a number of polygons on a display panel, which varied in size, shape, and arrangement. They were then required to select the matching binary code from two different options. In the next phase of training, the animals were shown a number of polygons and were required to "write" the matching binary number by lighting up one to three bulbs, with a lit bulb signifying "1" and an unlit bulb signifying "0" within the binary code. The animals learned this task, but it required hundreds of thousands of training trials.

Matsuzawa (1985) performed a more recent study of numerical matching with a chimpanzee named Ai. The chimpanzee was presented with groups of 3-dimensional objects, such as pencils and keys, and a keyboard with symbols representing specific objects, colors, and numbers. The goal of the task was to press the three keys that represented the identity, color, and number of items in the display. She succeeded in learning numerical labels for arrays of 1 to 6 items and correctly labeled 300 types of samples. Although Ai could press the keys in any order, she usually selected the color first, the identity second, and the number last, which may indicate that she was least confident about the numerical labels. In support of that theory, Matsuzawa noted that accuracy was lower for the numerical labels compared to the color and object identity labels and numerical labels were learned more slowly. In subsequent studies Ai was trained to label sets of heterogeneous objects and sets of dots that varied in size, density, and pattern with the Arabic numerals 1 through 7 (Murofushi, 1997) and to use the numeral 0 to represent the absence of items (Biro & Matsuzawa, 2001).

Studies also show that several bird species are capable of matching symbols to specific numerosities. For example, Xia et al. (2001) presented pigeons with exemplars of different numerosities, which were controlled for nonnumerical features. Each exemplar was followed by an array of letters and each letter was designated to correspond to one of the numerosities. With training, five pigeons learned to respond to the numerosities 1 through 4 at levels above chance, and two pigeons also learned to respond correctly to the numerosity 5.

In several studies, a grey parrot named Alex demonstrated the ability to respond to different quantities of items with a verbal numerical label (i.e., the word “three”) and to choose the matching quantity after hearing a verbal label (Pepperberg, 1994; Pepperberg & Gordon, 2005). For instance, when presented with a group of 4 corks he responded by verbalizing the word “four.” When presented with a collection of heterogeneous objects including 5 keys and asked, “What five?” he was able correctly to answer “keys.” These studies indicate that Alex could produce and comprehend verbal labels for quantities up to 6 and that his accuracy was unaffected by the size of the quantity. These data are interesting given the large number of animal numerical studies which show a decrease in accuracy as quantity increases.

In a subsequent study, the same parrot was able to combine two nonvisible sets of items and respond with the correct verbal label. A trial began when an experimenter placed two sets of variously sized objects on a tray and covered them with cups. The experimenter then briefly showed the parrot the items under each cup before covering them again. The cups were lifted one at a time so the parrot never saw all the items at once. The parrot was able to produce the correct verbal label for the total at levels greater than chance for the quantities 1 through 6 and his accuracy did not decrease as the total numbers of items increased. These results indicate that the parrot was able to remember the quantity under each cup, combine the quantities through some process, and then produce a label for the total amount (Pepperberg, 2006)

Absolute numerical judgments with analog quantities

Other researchers investigate absolute numerical knowledge in animals using a paradigm that requires the subject to select a fixed analog quantity from an array of simultaneously presented options. In one of the first known studies of absolute numerical ability in animals, a German scientist named Koehler (1950) trained a raven (*Corvus corax*) to choose a pot with 5 dots on the lid from among a set of pot lids depicting different numbers of dots. The raven could make this discrimination despite variations in the size and form of the dots. In another of Koehler's studies, a raven and a parrot were given a sample stimulus consisting of a lid bearing 1 to 6 dots and were required to match that sample to another lid bearing the same number of dots. The size and pattern of the dots on the sample and comparison lids were made as dissimilar as possible to prevent the birds from using perceptual cues instead of numerical cues to perform the task. Davis and Memmott (1992) warned, however, that the performance of Koehler's birds should be viewed with caution because his written accounts are lacking in experimental details.

Hayes and Nissen (1971) performed a very similar experiment with a chimpanzee named Viki and demonstrated that she could match a stimulus card displaying a random pattern of dots with a response card displaying the same number of dots. Performance was very good for dot patterns ranging in number from 1 to 3, but deteriorated as the number increased.

Other researchers found that rhesus monkeys (Hicks, 1956) and a raccoon (*Procyon lotor*; Davis, 1984) could choose a display of three items from displays consisting of 1 to 5 items, even when extraneous cues such as area, density, spatial arrangement, and odor were controlled. In addition, a recent laboratory study employing strict controls of nonnumerical factors demonstrated that monkeys were able to choose an exact numerical match for computer-generated examples of the numerosities 1 through 9, although performance was modulated by the ratio between the correct numerical match and the distractor stimulus (Jordan & Brannon, 2006).

The fact that animals can perform tasks requiring them to choose a specific numerosity from an array of alternatives or match a specific numerosity to a symbol or verbal label indicates that they have a sense of cardinality. That is, they have knowledge of the absolute value of the quantities. However, the results of these studies do not provide evidence that these animals can order the quantities, or symbols representing the quantities, along a continuum. In the study by Murofushi (1997), Ai learned to associate Arabic numerals with different quantities, but she may have represented Arabic numerals as separate and unrelated categories, just as we represent categories such as "tree" and "flower." There is some evidence that Ai failed to grasp the ordinal properties of the numerals. Even after extensive experience with the task, each new numeral still required the same number of trials to reach criterion as the numeral before. Ai did not recognize that each new numeral represented the addition of one more item to the sample array, which would have allowed for immediate transfer to new numerals. In this way, Ai is different from preschool children who immediately recognize that a new numeral is connected to a new quantity and that each new numeral in a series represents one more item than the one before (Carey, 2004). Despite Ai's inability to demonstrate ordinal knowledge in this particular task, numerous studies that do reveal ordinal knowledge in animals have already been discussed in this chapter (e.g., Brannon & Terrace, 1998, 2000; Emmerton et al., 1997; Judge et al., 2005; Kilian et al., 2003; Olthof et al., 1997; Washburn & Rumbaugh, 1991). Collectively, these findings suggest that both ordinal and cardinal studies are necessary to develop a true picture of animal numerical abilities.

Numerical discriminations of sequentially presented objects or events

All of the studies described thus far have involved simultaneously presented stimuli, but another important aspect of numerical cognition is the ability to keep track of sequentially presented stimuli or events. Adult humans have a concept of number that allows them to keep track of a sequential number of lightening flashes, or trips to the store, or notes of music, but whether animals share this ability is an interesting experimental question. It is plain to see how the ability to compare two visible quantities of food or to weigh a number of attackers against a number of allies would be beneficial to animals in the wild, but the benefits of keeping track of sequentially encountered objects or events is less obvious. Sequential studies are also different than simultaneous studies because different methods of control must be used to rule out nonnumerical factors. Instead of controlling for stimulus features such as size and density, researchers must control for temporal factors such as total duration of presentation and rate of presentation. In recent years, researchers using a variety of paradigms and methods of control have provided a wealth of evidence for sequential numerical abilities in nonhuman animals.

In a study by Beran (2001), two chimpanzees watched as pieces of candy were placed, one at a time, into two opaque cups. The cups were lined with foam to reduce auditory feedback and temporal factors were controlled. The chimpanzees succeeded in choosing the cup with the largest quantity when sets of 1 to 9 candies were used and the maximum difference did not exceed 4 candies. They also succeeded when the quantities in each cup were presented as two smaller sets of 1 to 6 candies and even when they were presented as three smaller sets of 1 to 4 candies. These results provide evidence that chimpanzees can mentally combine and compare sequentially presented items, but the candies were uniform in size so it is unclear whether the animals were using number or amount of food. Accuracy in all of these experiments was significantly correlated with the ratio and total number of the two quantities. In the final experiment, 1 to 5 candies were placed into each cup and the experimenter then removed one of the candies from the cup on the left. One of the chimpanzees was able to select the cup with the largest amount of candy on these trials, which indicates that he was representing the absolute difference between the two cups and not just keeping track of which cup had the larger amount. It is possible that he was representing the absolute magnitude of each quantity as well, but it is not clear from this procedure.

In a variation of this study, chimpanzees watched as two sets of 1 to 10 marshmallows were sequentially placed into opaque cups. After these sequential presentations, a visible set was made available as well and the chimpanzees were allowed to select any of the three choices. The animals succeeded in selecting the greatest quantity, which indicates that they remembered the approximate magnitude of the largest nonvisible set and were able to compare that to the visible set. The chimpanzees also responded correctly when one marshmallow was removed from one of the two sequentially presented sets before selection, but not when more than one marshmallow was removed (Beran, 2004). Similar studies have shown that orangutans are able to compare sequentially presented sets of food items and represent the absolute difference between them (Call, 2000) and that rhesus monkeys are able to choose the larger of two sequentially presented computer-generated sets of items, even when number does not covary with quantity (Beran, 2007).

Recent research has shown that capuchin monkeys, like chimpanzees and other species more closely related to humans, are also sensitive to the quantitative value of sequentially presented stimuli. Evans, Beran, Harris, and Rice (2008) presented capuchin monkeys with one set of simultaneously visible food items and a second set of sequentially presented items that was never visible in its entirety. Results revealed that the monkeys exhibited high accuracy in choosing the larger set, regardless of whether the correct set was the simultaneously visible or sequentially presented set. However, the monkeys exhibited near-chance performance in a second experiment in which they were required to choose between two sequentially presented sets. Further analysis of the results revealed that performance was related to trial duration, which suggests that their poor performance may have stemmed from a relatively limited attentional capacity.

Lewis et al. (2005) used a search task to investigate the ability of lemurs (*Eulemur mongoz*) to keep track of sequentially presented quantities. The animals watched as an experimenter placed grapes, one at a time, into an opaque bucket. The lemurs were then allowed to retrieve the grapes from the bucket. On half of the trials, grapes were placed into a false bottom in the bucket so that they were inaccessible to the animals. The experimental question was whether or not the lemurs would spend more time searching the bucket when grapes should have remained compared to when all the grapes were retrieved. The lemurs searched longer on trials with hidden grapes when the numerosities differed by a 1:2 ratio, but not when they differed by a 2:3 ratio or a 3:4 ratio. In a control condition, the lemurs watched as two half-grapes were placed into the bucket and then continued to search after retrieving a whole grape. This indicates that they were representing the number of grapes and not just the total amount of grape.

There are several other paradigms that researchers use to investigate number representations in sequential studies involving animals. Davis and Bradford (1986) required rats to enter a particular tunnel (e.g., the third tunnel) in an array of six tunnels to obtain food. Spatial, olfactory, and visual cues were controlled to ensure that the rats were using the number of sequentially encountered tunnels as a cue to the correct tunnel. The rats even performed well when they were presented with a novel configuration that required them to turn a corner in order to select the correct tunnel. In an extension of this study, rats were able to perform above chance when the correct tunnel was the tenth among twelve or eighteen tunnels and when the size of the tunnels was varied to control for cumulative length (Suzuki & Kobayashi, 2000). Other researchers used a similar procedure to show that 5-day-old chickens (*Gallus gallus*) could choose the correct food well from an array of identical food wells using numerical cues (Rugani, Regolin, & Vallortigara, 2007).

Other investigators have used a forced-choice discrimination procedure to study sequential numerical knowledge. Keen and Machado (1999) presented pigeons with two different flashing lights and the pigeons learned to choose the light that flashed the least number of times. In another study, Roberts and Mitchell (1994) trained pigeons to discriminate between two and eight flashes of light. Davis and Albert (1986) conducted a sequential study using auditory stimuli. In this study, rats were rewarded for lever pressing after hearing three noise bursts and not after hearing two or four noise bursts. The rats learned to lever press more after three noise bursts than two or four, which indicates that responding was based on something more complicated than a "same" versus "many" discrimination.

Numerical discriminations of sequential motor responses

In other sequential paradigms, the numerical stimuli to be discriminated consist of the animal's own responses. Procedures requiring the subject to make a discrete number of motor responses are sometimes referred to as constructive enumeration tasks. These tasks contrast with receptive enumeration tasks, in which the animal must enumerate externally presented sets of items or events before making a response (Beran & Rumbaugh, 2001; Xia et al., 2000). In a paradigm developed by Mechner (1958), rats were required to make a certain number of responses in order to obtain a reward. After 4, 8, 12, or 16 presses on lever A the rats could switch to lever B to obtain food. Switching levers prematurely was an error and resulted in the trial starting over again. In a subsequent study, Platt and Johnson (1971) required rats to signal when they had completed a certain number of lever presses by poking their nose into a food tray with a sensor. In both of these studies the number of presses on the first lever was approximately normally distributed around a number slightly higher than the number of required presses. This is a rational strategy given that abandoning the lever too early resulted in a penalty. Even after training, results were imprecise and revealed scalar variance, which is a hallmark of the analog magnitude model. This lever-pressing paradigm allowed the number of responses to covary with the time spent responding, but a later experiment by Mechner and Guevrekian (1962) manipulated the rate of responding by changing the level of the rat's food deprivation and found that this manipulation had no effect on the average number of presses or the distribution.

In a unique study of sequential enumeration ability, Boysen and Berntson (1989) baited several sites around their laboratory with 1 to 3 oranges each. A chimpanzee named Sheba learned to move from site to site and then choose the Arabic numeral that corresponded to the total number of oranges she had encountered. In subsequent testing, the researchers replaced the oranges with cards depicting Arabic numerals ranging from 0 to 4 and her performance was not disrupted. Sheba spontaneously summed the Arabic numerals at three different sites and chose the numeral that represented the total value. It is possible that Sheba was summing the quantities represented by each individual numeral, but it is also possible that she was using a simpler strategy such as "counting all" or "counting on." In the counting all strategy, the subject enumerates all the items sequentially and maintains a running tally. In the counting on strategy, the subject begins with the cardinal value of the first quantity and then adds onto it by enumerating the rest of the items sequentially. Boysen (1993) referred to these two strategies as "pseudo-addition" because true addition involves combining a cardinal representation of one numerosity with the cardinal representation of another numerosity.

Xia et al. (2000) investigated the use of symbols in a sequential task by presenting pigeons with a key that randomly displayed one of several possible symbols on each trial. The symbols indicated how many pecks the pigeons were required to make to the key. After completing a series of pecks, the pigeons signaled the completion of the trial by pecking a second key. Six pigeons were able to match 1 to 5 pecks with the corresponding symbols and four were also able to match 6 pecks to the corresponding symbol. Temporal variables were not controlled during this study, but subsequent analysis indicated that the animals were mainly relying on number and not time. As with most animal numerical studies, accuracy decreased as the number of required responses increased and the distribution of errors increased systematically with the number of responses. This is a more flexible form of constructive enumeration than that demonstrated by the rats in Mechner's (1958) paradigm because the rats were trained with each number successively. The use of symbols as a cue to the number of required pecks allowed researchers to present the pigeons with randomly intermixed series requiring different numbers of pecks on every trial.

In a series of constructive enumeration studies, two chimpanzees named Lana and Austin used a joystick to contact the number of boxes on a computer screen that equaled a randomly selected Arabic numeral (1-3 for Lana and 1-4 for Austin). After they finished contacting boxes, they made contact with the numeral to signal the end of the trial. No visual feedback was provided so the chimpanzees had to rely on their memory to determine whether they had removed the correct number of boxes. They did not use any specific pattern of selection, which suggests that they knew that the number of items was the important factor and not the selection pattern (Beran et al., 1998; Rumbaugh et al., 1989; Rumbaugh & Washburn, 1993).

In a more recent study using a similar procedure, Beran and Rumbaugh (2001) provided evidence that one chimpanzee could contact a number of dots equal to the Arabic numerals 1 through 6 and a second chimpanzee could also respond correctly to the numeral 7. There were a maximum of ten dots present on every trial and they were randomly positioned so that every trial had a unique pattern. Analysis of trial duration data indicated that the chimpanzees were responding based on number and not temporal cues and analysis of accuracy indicated that performance decreased as the target number increased.

In another study involving successive actions, Capaldi and Miller (1988) presented rats with series of reinforced (R) and nonreinforced (N) runway trials. The rats received multiple presentations of RRN and NRRN series and developed a pattern of running more slowly on the terminal N trials of each series than the other reinforced trials. The researchers concluded that the rats were using the number of completed R trials to predict when the N trials would occur for each series. Davis and Pérusse (1988) argued that the rats could have been using the rhythmic pattern of events, rather than the number of events, to predict when the N trial would occur. To test this hypothesis, Burns et al. (1995) performed a similar runway experiment with rats in which they systematically varied the inter-trial intervals from 20 to 120 seconds, which would have disrupted any temporal rhythms. The rats quickly developed a pattern of running slowly on the terminal N trials, which suggests that they were using number as a cue.

Rats have also been trained to discriminate between different series of reinforced and nonreinforced trials using brightness and texture cues on the runway floor. For instance, Burns, Dunkman, and Detloff (1999) consistently presented rats with a rough and white floor during an XNY series (where the X and Y represented different food items) and a smooth and black floor during a ZNN series (where the Z represented a third type of food item). Using this procedure, the researchers were able to compare performance between more than one series in the same group of rats. For both series, the rats developed faster running for rewarded trials than for nonrewarded trials.

Monkeys previously trained to make ordinal judgments using Arabic numerals provide a unique opportunity to study the use of numerical cues and spontaneous transfer between series. Arabic numerals, instead of the texture of runway floors, can be used as a cue to help the monkeys determine which type of series is being presented. This in turn, could act as a cue to help them predict when a nonreinforced trial will occur. In **Chapter 4: Macaques (*Macaca mulatta*) use of numerical cues in maze trials**, we trained four Arabic numeral-experienced monkeys on an RRRN series. The goal of the maze was an Arabic numeral 3, which corresponded to the number of reinforced maze trials in the series. We then introduced probe series involving different numbers of reinforced trials. Two of the monkeys were given probe series of the numerals 2 and 4, intermixed with the familiar 3 series, and the remaining two monkeys were given probe series of the numerals 2 through 8. As was true during training, the Arabic numeral displayed in the maze corresponded to the number of reinforced trials that would occur before one nonreinforced trial. We hypothesized that the monkeys would use the numerals to discriminate among different series of reinforced and nonreinforced trials and anticipate the nonreinforced trials. We also hypothesized that the monkeys' prior knowledge of Arabic numerals would allow for spontaneous transfer from one Arabic numeral to another during this sequential task.

During training on the RRRN series, two of the four monkeys developed a "slow, fast, faster, slow" pattern, which suggested they were anticipating the final nonreinforced trial. The other two monkeys performed gradually slower on each trial in a series, which made it impossible to ascertain whether or not they were predicting precisely when the final trial would occur. During testing, the monkeys receiving probe trials of the numerals 2 and 4 showed some generalization to the new numerals and developed a pattern of performing more slowly on the nonreinforced trials than the reinforced trials, indicating the use of the changing target numeral cues to anticipate those final trials. The monkeys receiving probe trials of the numerals 2 through 8 did not use the changing numeral to predict precisely when the nonreinforced trial would occur in each series, but they did incorporate the changing numerals into their strategy by performing faster overall on series with greater target numerals.

This study provided evidence that number-trained rhesus monkeys could use Arabic numerals as a cue to facilitate performance on a task involving sequential responses. However, the pattern established by two of the monkeys during training of performing gradually slower on each trial in a series, and the failure of the monkeys receiving probe trials of the numerals 2 through 8 to generalize the pattern learned during training to new target numerals, highlighted the need for a task that specifically addressed the monkeys' understanding of the number of trials in a series.

In Experiment 2 of **Chapter 2: What do Arabic numerals mean to macaques?**, the monkeys were provided with two Arabic numeral cues in a computerized maze and each numeral was "baited" with the corresponding number of pellets. Moving the cursor into contact with either numeral resulted in the delivery of a pellet, unless the monkey had already earned the corresponding number of pellets for that problem (e.g., the numeral 4 would only be reinforced four times in a problem). We reasoned that a monkey could travel to the larger numeral the corresponding number of times and then behaviorally indicate that he knew he had exhausted the pellets at that location by traveling to the smaller numeral. In contrast, if the monkeys know only the ordinal and not absolute values corresponding to the numerals, then they would have no basis for knowing when to stop responding to the larger of the two numerals. This design allowed us to assess the monkeys' understanding of the cardinal value of the numerals in a sequential task.

In **Chapter 5: Rhesus monkeys (*Macaca mulatta*) select Arabic numerals or visual quantities corresponding to a number of sequentially completed maze trials**, we further investigated the symbolic nature of Arabic numerals in a sequential enumeration task. We hypothesized that the monkeys in the Chapter 4 study did not use the target numbers in the way anticipated because there was little motivation for the monkeys to keep track of the absolute number of trials. The reinforcement pattern remained the same regardless of the strategy used by the monkeys to perform the task. During training, for instance, the monkeys always received three reinforced trials followed by one nonreinforced trial, regardless of how quickly they completed each maze trial. In addition, the monkeys performed thousands of trials a day on this and other tasks, so a few nonreinforced trials were probably not very salient.

In Chapter 5, four number-trained rhesus monkeys were trained to enumerate their sequential responses and reinforced only when they had made a correct response. This should increase motivation to perform at high levels because of the time invested in each series. After completing a series of computerized maze trials, the monkeys were given a same/different discrimination involving a numerical stimulus (an Arabic numeral or a dot array) and the letter D (for “different”). The goal was to choose the numerical stimulus if it matched the number of just-completed maze trials, and to choose the D if it did not. Previous studies have shown that nonhuman primates are capable of representing, combining, and comparing nonvisible, sequentially presented sets of items (e.g., Beran, 2001; Call, 2000; Hauser et al., 2000), but this study tested the ability of monkeys to enumerate their own responses and match the number of responses with the corresponding Arabic numeral or visual array.

Numerical discriminations of sequentially presented auditory and tactile stimuli

The stimuli used in numerical tasks are most often visual, but auditory stimuli also can be used in sequential tasks. In a recent study, Hauser, Dehaene, Dehaene-Lambertz, and Patalano (2002) used a habituation-dishabituation paradigm to investigate auditory numerical discrimination in cotton-top tamarins (*Saguinus oedipus*). In the habituation phase, some monkeys were presented with sequences of two speech syllables and others were presented with sequences of three speech syllables. The syllables varied in overall duration, inter-syllable duration, and pitch. In the test phase, all of the monkeys were presented with counterbalanced sequences of either two or three tones that also varied in overall duration, inter-tone duration, and pitch. Results revealed that the monkeys looked reliably longer at test stimuli that differed in number from the stimuli with which they were trained. These results indicate that monkeys can represent the numerical value of auditory stimuli and that their representations are abstract enough to accommodate differences in format.

There is also evidence that animals can perform numerical discriminations based on successive tactile stimuli. Davis, MacKenzie, and Morrison (1989) stroked the whiskers of a group of rats two, three, or four times in a row to indicate which arm of a Y-maze they should enter. The rats learned to enter the correct maze arm, which indicated that they were able to transfer sequential tactile numerical information to the visual maze task.

Spontaneous numerical operations in nonhuman animals and human infants

Researchers often use a sequential presentation method to investigate spontaneous addition and subtraction abilities in animals and human infants. In the expectancy-violation paradigm, which was originally developed for use with human infants, an untrained subject watches as items are successively placed behind an opaque screen. Typically, each subject is tested in only one condition. On some of the trials, items are covertly removed or added from behind the screen. The screen is then raised to reveal all of the items and looking time is recorded. According to this paradigm, unexpected outcomes should produce longer looking times in comparison to expected outcomes.

Hauser and Carey (2003) used the expectancy-violation paradigm to study representations of small numbers of objects in free-ranging rhesus monkeys. The monkeys watched as experimenters placed eggplants, one at a time, behind a screen. The screen was then raised to reveal a possible or impossible outcome. The monkeys exhibited longer looking times for the impossible outcomes when presented with 1 + 1 = 2 or 3, 2 + 1 = 2 or 3, and 2 + 1 = 3 or 4, but failed at 2 + 2 = 3 or 4, and problems consisting of more than two quantities of eggplants, such as 1 + 1 + 1 = 3. To test the possibility that the monkeys were representing continuous variables such as volume, contour length, or visible surface area instead of the number of objects, the experimenters placed two small eggplants behind the screen and then raised the screen to reveal two small eggplants or one big eggplant that was roughly twice the size of a small eggplant. The monkeys exhibited longer looking times for the impossible outcome of one big eggplant, which suggests that they were not relying solely on volume or surface area cues. Uller, Hauser, and Carey (2001) provided additional evidence that monkeys do not rely on continuous perceptual variables in the expectancy-violation paradigm. They found that cotton-top tamarins exhibited longer looking times for the outcome that was a numerical mismatch in a 1 small + 1 small = 2 small objects or 1 big object problem, where the single large outcome matched the expected outcome in volume and surface area.

In a similar study conducted by Flombaum et al. (2005), rhesus monkeys watched as lemons were placed behind a screen. To rule out continuous variables as a possible cue, the amount of lemon was equated in the two conditions by using lemons that were larger or smaller in size than the lemons placed behind the screen. The monkeys exhibited longer looking times for numerical violations than expected numerical outcomes. None of the lemons that emerged from behind the screen were identical to the lemons placed behind the screen, which suggests that the monkeys were using the number of items as a cue rather than the physical identity of the items. Interestingly, the monkeys recognized that 4 + 4 = 8 rather than 4, but they showed no difference in looking time when they were shown 2 + 2 and tested with an outcome of 4 or 6. These results indicate that the monkeys were able to discriminate large sets of items, but only when the ratio between the observed and expected outcome was small. Other animals, such as lemurs and dogs (*Canis familiaris*), also show numerical expectations when given simple tests, such as 1 + 1 = 2 or 3 (Santos, Barnes, & Mahajan, 2005; West & Young, 2002).

In a study designed to assess spontaneous subtraction in nonhuman primates using a different paradigm, untrained rhesus monkeys were presented with a quantity of plums on one stage and a second quantity on another stage. The experimenter subsequently occluded both stages and removed one or no plums from each stage. The monkeys were then allowed to approach a stage and eat the plums behind the occluder. The monkeys successfully chose the larger quantity of plums following the subtraction of one piece of food from two or three pieces of food. Accuracy was high regardless of whether food was subtracted from one or both of the initial quantities (Sulkowski & Hauser, 2001).

The results of these spontaneous numerical operation studies with animals closely mirror the results observed in human infants. For example, Wynn (1992) conducted a study in which 5-month-old infants watched an experimenter placed two Mickey Mouse dolls behind a screen. When the screen was removed, the infants exhibited longer looking times for the unexpected outcome of one doll or three dolls than they exhibited for the expected outcome of two dolls. Similar results were obtained for subtraction problems in which the experimenter placed two dolls behind a screen and then removed one doll. The infants looked longer at the impossible outcome of two or zero dolls than the expected outcome of one doll.

There is conflicting evidence regarding the use of continuous variables by infants in these types of studies. In a study by Feigenson, Carey, and Spelke (2002) 7-month-old infants watched as experimenters placed two small objects behind a screen. The screen was then removed to reveal one big object (which had the expected surface area and volume, but an unexpected number) or two big objects (which had an unexpected surface area and volume, but the expected number). The infants looked longer at the unexpected surface area and volume outcome, but not at the unexpected number outcome, which suggests that continuous variables were underlying their representations. In contrast to this finding, Uller (1997) found that 8-month-old infants exhibited longer looking times for the outcome that was a numerical mismatch in a 1 small + 1 small = 2 small or 1 big comparison, where the big object matched the expected outcome in volume and surface area.

These expectancy-violation studies have been used as evidence that human infants and some nonhuman animal species have an innate understanding of simple arithmetic operations. In contrast to this view, Simon (1997) suggested that the animals in these studies were not performing arithmetic, but instead were tracking each item placed behind the screen. According to this view, the object tracking system assigns a unique index for each object placed behind the screen and when the screen is removed the indexes are placed in one-to-one correspondence with the revealed set. Looking time is increased not because the subject recognizes the outcome as numerically incorrect, but because the subject detects a mismatch between the number of visible items and the number of indexes. Subsequent findings suggest that these representations may not contain object identity or location information. Simon, Hespos, and Rochat (1995) found that the expectations of infants were not violated when Elmo dolls were placed behind the screen and then secretly replaced with the correct number of Ernie dolls. Koechlin, Dehaene, and

Mehler (1998) placed dolls on a revolving surface and showed that the expectations of infants were not violated due to a change in location of the objects.

Numerical discriminations in human infants

In addition to the expectancy-violation paradigm, several other paradigms have been used to study numerical capacities in infants and young children. Wynn (1996) used a habituation-dishabituation paradigm to study the ability of 6-month-old infants to discriminate different numbers of visual events. The visual events were "puppet jumps" created by an experimenter moving a toy puppet up and down. Half of the infants were habituated to two puppet jumps and the other half were habituated to three puppet jumps. Then, both groups of infants were tested with series of two and three jumps. The timing of the jumps was carefully controlled so that the duration was not a clue. The infants looked longer at the novel number of jumps, which indicated that they could discriminate one series of jumps from the other.

Other studies using the habituation-dishabituation paradigm have shown that infants in the first year of life can discriminate between sets of simultaneously presented objects on the basis of numerosity (Antell & Keating, 1983; Starkey & Cooper, 1980; Strauss & Curtis, 1981). In these studies, infants as young as 4-days-old discriminated 1 versus 2 objects and 2 versus 3 objects, but failed at comparisons of 3 versus 4 and 3 versus 5 objects. .

Feigenson and Carey (2003, 2005) used search behavior as a measure of numerical ability. Children 12 to 14 months old watched as an experimenter placed a number of objects (balls or crackers) into a box and were then allowed to retrieve the objects. On some trials, experimenters secretly placed one or more of the objects into a hidden compartment at the bottom of the box so that they were inaccessible to the children. The children searched significantly longer after retrieving 1 object from the box when they had seen 2 objects being placed in the box compared to 1 object. They also searched longer after retrieving 2 objects when they had observed 3 objects being placed in the box compared to 2 objects. However, when infants viewed 4 objects being placed into the box they did not search longer after retrieving 2 objects than infants who had seen 2 objects being placed into the box. The authors argued that the infants failed to represent arrays of more than 3 objects.

Bijeljac-Babic, Bertoncini, and Mehler (1993) used sucking rhythm to study numerical abilities in young infants. The infants were allowed to suck on a rubber nipple connected to a pressure gauge and computer. Whenever the infant sucked on the nipple the computer delivered a nonsense word, such as “bafikoo,” through a loudspeaker. The duration of the words and rate of speech were highly variable, but the number of syllables remained constant. When the infant habituated to the number of syllables, the rate of sucking dropped and the computer switched to a different number of syllables. Results revealed that a switch in the number of syllables was accompanied by renewed vigor in sucking behavior. A control group in which novel words were introduced with no change in the number of syllables showed no reaction.

Although these studies provide evidence that infants can discriminate between different number categories, they do not address whether or not infants understand the order of these categories. Brannon (2002) addressed the issue of ordinal knowledge in a study with 11-month-old infants. At the start of the experiment, the infants were habituated to sequences of dot arrays on a computer screen that increased or decreased (e.g., 4-8-16 or 16-8-4). Infants were then tested with new numerical values that alternated between increasing and decreasing sequences. The infants trained with increasing ordinal sequences looked longer at decreasing ordinal sequences and vice versa, which indicates that they understood the ordinal relations among the number of dots. Infants who were 9 months old failed to discriminate between the increasing and decreasing sequences. Thus, ordinal knowledge may develop later than the ability to discriminate stimuli based on number.

Cross-modal and cross-procedural transfer tests

Another paradigm used to investigate nonverbal numerical ability involves transfer tests that require an animal or human infant to generalize numerical knowledge across modalities or procedures. These tests provide evidence pertaining to the controversy of whether or not human infants and nonhuman animals represent number abstractly. In other words, whether these nonverbal populations understand that sets of stimuli differing in perceptual features and modality, such as three visible squares, three tones, and three flashes of lightening share the cardinal value three. Gelman and Gallistel (1978) recognized abstractness as one of the five principles of formal counting and Davis and Pérusse (1988) further argued that the ability to abstract number across different contexts and modalities is necessary for a true concept of number. Evidence from functional imaging studies suggests that the posterior parietal may play a role in abstract number processing because this area is activated by numerical stimuli in humans and rhesus monkeys, regardless of modality (Nieder, Diester, & Tudusciuc, 2006).

Little research has been conducted to determine whether infants and animals have an abstract representation of number equal to that of adult humans. This lack of evidence led Dehaene (1997) to conclude that the brains of animals and human babies are not as flexible as adult human brains and that those rigid brains, “work their minor arithmetical miracles only within quite limited contexts” (p. 5). The research that has been conducted tends to focus on transfer of numerical knowledge across nonnumerical perceptual features such as size and color, across modalities such as auditory or visual, and across sequential and simultaneous presentation methods.

In studies by Starkey and colleagues (Starkey, Spelke, & Gelman, 1983, 1990) infants were able to detect numerical correspondences between the visual and auditory modalities. Infants 6 to 9 months old were presented with visual displays containing two and three items while listening to two or three drumbeats. Researchers found that when the infants were listening to three drumbeats they reliably looked at the visual display of three objects. When hearing to two drumbeats they reliably looked at the display of two objects. In similar studies, Jordan and colleagues (Jordan, Brannon, & Gallistel, 2006; Jordan, Brannon, Logothetis, & Ghazanfar, 2005) demonstrated that rhesus monkeys and 7-month-old infants preferred to look at video-clips containing a number of conspecifics equal to the number of vocalizations they heard. These studies suggest that rhesus monkeys and human infants possess an abstract concept of number that reaches across two sensory modalities.

In another study, Fernandes and Church (1982) presented rats with sequences of white noise and rewarded them for pressing a lever on the right when they heard two noise bursts and a lever on the left when they heard four. The rats learned to respond based on the number of noise bursts, even when temporal cues were controlled by varying the duration of each burst, as well as the total duration of the auditory sequences. When the experimenters substituted light flashes for sounds, the rats immediately transferred their knowledge to the new task, which suggests that their representation of number was not tied to the auditory modality.

In a similar study, Church and Meck (1984) taught rats to press a lever on the left after viewing a sequence of two flashes or hearing a sequence of two white noise bursts, and a lever on the right after viewing a sequence of four flashes or hearing a sequence of four noise bursts. When the rats were then presented with a combination of two lights and two noise bursts they spontaneously integrated the number of visual and auditory stimuli and responded by pressing the right lever. This outcome indicates that the rats based their behavior on an abstract, amodal representation of number.

Davis and Albert (1987) trained rats on a more complex task that required them to discriminate between two, three, or four bursts of noise. When the experimenters substituted light flashes for the noise bursts they found no evidence of transfer. These results, combined with the results of the Fernandes and Church (1982) study, suggest that abstract representations in rats may be confined to simple tasks requiring only a “few” and “many” judgment.

The monkeys in our laboratory have had extensive experience with many types of numerical tasks, but it is unclear whether they possess an abstract numerical concept that allows transfer of numerical knowledge from one type of task to another. The goal of **Chapter 6: Numerical abstraction across presentation mode by rhesus monkeys** was to investigate whether the monkeys in our laboratory could transfer learning in a sequential numerical task to a simultaneous numerical task. During training, the monkeys learned to make one response after viewing a sequence of three circles flashed on a computer screen and another response after viewing a sequence of seven circles flashed on a computer screen. The monkeys were then presented with nonreinforced probe trials consisting of groups of three or seven simultaneously visible circles. The goal was to assess whether or not the monkeys would transfer the numerical knowledge gained in the sequential task to the simultaneous task by spontaneously providing a “three” response when presented with three simultaneously visible circles and a “seven” response when presented with seven simultaneously visible circles. Evidence of transfer would suggest that the monkeys possess an abstract representation of number that is not tied to a specific mode of presentation.

In a second experiment, a variation of the standard transfer paradigm was employed to investigate abstract number concept in monkeys. The same monkeys from the first experiment were used in this experiment so they all had experience making one response after viewing a sequence of seven circles and another response after viewing a sequence of three circles. In this second experiment the monkeys were trained on a new task that involved groups of three or seven simultaneously visible circles. For half of the monkeys, the correct response when presented with three simultaneously visible circles was the same as the correct response when presented with three sequential circles in the first part of the study. Similarly, the correct response when presented with seven simultaneous circles was the same as the correct response when presented with seven sequential circles. For the other half of the monkeys, the reward contingencies were reversed so that the correct response when presented with three simultaneously visible circles was the same as the correct response for seven sequential circles and the correct response for seven simultaneous circles was the same as the correct response for three sequential circles. If the monkeys categorized sequentially and simultaneously presented stimuli together on the basis of number then it should take the group with reversed reward contingencies longer to learn this task than the group for which the reward contingencies stayed the same. Thus, the results of this study help to answer the question of whether or not the monkeys in our laboratory, like humans, have an abstract concept of number that spans different contexts and methodologies.

Overview of dissertation

The subsequent chapters in this dissertation consist of a variety of control and transfer studies designed to illuminate what Arabic numerals symbolize to rhesus monkeys. In Chapters 2 and 3, the monkeys were required to compare and order Arabic numerals and were rewarded with either proportional or probabilistic rewards. These studies provided information on whether numerical discriminations are based on the numerousness attribute of the stimuli or nonnumerical attributes such as hedonic value and conditioned 2-choice discriminations. They also provided information on whether numerals symbolize absolute or ordinal knowledge to the monkeys. In chapters 4 and 5, the monkeys were required to enumerate their own sequential responses and associate that quantity with an Arabic numeral. These studies provided data regarding the use of absolute versus ordinal knowledge by monkeys in sequential tasks. Data from all of these studies were examined to determine if representations were approximate or inexact, which provided information on the underlying mental mechanisms. The study in Chapter 6 was designed to investigate the generality of the monkeys' symbolic number concept using transfer tests. The ability to abstract number across presentation mode would indicate that numerals are truly symbolic and not simply functioning as part of a specific stimulus-response-reward association. Taken together, these studies shed light on the nature of the monkeys' number concept and whether the animals' understanding of Arabic numerals is symbolic in the same way that it is for humans.

References

Almeida, A., Arantes, J., & Machado, A. (2007). Numerosity discrimination in preschool children. *Journal of the Experimental Analysis of Behavior, 88,* 339-354.

Anderson, U. S., Stoinski, T. S., Bloomsmith, M. A., Marr, M. J., Smith, A. D., & Maple, T. L. (2005). Relative numerousness judgment and summation in young and old western lowland gorillas. *Journal of Comparative Psychology, 119,* 285-295.

Antell, S. E., & Keating, D. P. (1983). Perception of numerical invariance in neonates. *Child Development, 54,* 695-701.

Atkinson, J., Campbell, F. W., & Francis, M. R. (1976). The magic number 4 +/- 0: A new look at visual numerosity judgments. *Perception, 5,* 327-334.

Barth, H., La Mont, K., Lipton, J., Dehaene, S., Kanwisher, N., & Spelke, E. (2006). Non-symbolic arithmetic in adults and young children. *Cognition, 98,* 199-222.

Bell, E. T. (1945). *The development of mathematics*. New York: McGraw-Hill.

Beran, M. J. (2001). Summation and numerousness judgments of sequentially presented sets of items by chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology, 115,* 181-191.

Beran, M. J. (2004). Chimpanzees (*Pan troglodytes*) respond to non-visible sets after one-by-one addition and removal of items. *Journal of Comparative Psychology, 118,* 25-36.

Beran, M. J. (2007). Rhesus monkeys (*Macaca mulatta*) enumerate large and small sequentially presented sets of items using analog numerical representations. *Journal of Experimental Psychology: Animal Behavior Processes, 33,* 42-54.

Beran, M. J., Beran, M. M., Harris, E. H., & Washburn, D. A. (2005). Ordinal judgments and summation of nonvisible sets of food items by two chimpanzees (*Pan troglodytes*) and a rhesus macaque (*Macaca mulatta*). *Journal of Experimental Psychology: Animal Behavior Processes, 31,* 351-362.

Beran, M. J., Evans, T. A., & Harris, E. H**.** (2008). *Perception of food amounts by chimpanzees based on the number, size, contour length and visibility of items*. Manuscript submitted for publication.

Beran, M. J., Evans, T. A., Leighty, K. A., Harris, E. H., & Rice, D. (2008). Summation and quantity judgments of sequentially presented sets by capuchin monkeys (*Cebus apella*). *American Journal of Primatology, 70,* 191-194.

Beran, M. J., Harris, E. H., Evans, T. A., Klein, E. D., Chan, B., Flemming, T. M., & Washburn, D. A. (2008). Ordinal judgments of symbolic stimuli by capuchin monkeys (*Cebus apella*) and rhesus monkeys (*Macaca mulatta*): The effects of differential and nondifferential reward. *Journal of Comparative Psychology, 122,* 52-61.

Beran, M. J., & Rumbaugh, D. M. (2001). “Constructive” enumeration by chimpanzees (*Pan troglodytes*) on a computerized task. *Animal Cognition, 4,* 81-89.

Beran, M. J., Rumbaugh, D. M., & Savage-Rumbaugh, E. S. (1998). Chimpanzee (*Pan troglodytes*) counting in a computerized testing paradigm. *Psychological Record, 48,* 3-19.

Beran, M. J., Taglialatela, L. A., Flemming, T. M., James, F. M., & Washburn, D. A. (2006). Nonverbal estimation during numerosity judgments by adult humans. *The Quarterly Journal of Experimental Psychology, 59,* 2065-2082.

Bijeljac-Babic, R., Bertoncini, J., & Mehler, J. (1993). How do 4-day-old infants categorize multisyllabic utterances? *Developmental Psychology, 29,* 711-721.

Biro, D., & Matsuzawa, T. (2001). Use of numerical symbols by the chimpanzee (*Pan* *troglodytes*): Cardinals, ordinals, and the introduction of zero. *Animal Cognition, 4,* 193-199.

Boysen, S. T. (1993). Counting in chimpanzees: Nonhuman principles and emergent properties of number. In S. T. Boysen & E. J. Capaldi (Eds.), *The development of numerical competence: Animal and human models* (pp. 39-59). Hillsdale, NJ: Lawrence Erlbaum Associates.

Boysen, S. T., & Berntson, G. G. (1989). Numerical competence in a chimpanzee (*Pan troglodytes*). *Journal of Comparative Psychology, 103,* 23-31.

Boysen, S. T., & Berntson, G. G. (1995). Responses to quantity: Perceptual versus cognitive mechanisms in chimpanzees (*Pan troglodytes*). *Journal of Experimental Psychology: Animal Behavior Processes, 21,* 82-86.

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,* 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: A Multidisciplinary Journal, 24,* 423-443.

Boysen, S. T., Mukobi, K. L., & Berntson, G. G. (1999). Overcoming response bias using symbolic representations of number by chimpanzees (*Pan troglodytes*). *Animal Learning and Behavior, 27,* 229-235.

Brannon, E. M. (2002). The development of ordinal numerical knowledge in infancy. *Cognition, 83,* 223-240.

Brannon, E. M., & Roitman, J. D. (2003). Nonverbal representations of time and number in animals and human infants. In W. H. Meck (Ed.), *Functional and neural mechanisms of interval timing* (pp. 143-182). Boca Raton, FL: CRC Press.

Brannon, E. M., & Terrace, H. S. (1998). Ordering of the numerosities 1–9 by monkeys. *Science, 282*, 746–749.

Brannon, E. M., & Terrace, H. S. (2000). Representation of the numerosities 1-9 by rhesus macaques (*Macaca mulatta*). *Journal of Experimental Psychology: Animal Behavior Processes, 26,* 31–49.

Brannon, E. M., & Terrace, H. S. (2002). The evolution and ontogeny of animal numerical ability. In M. Bekoff, C. Allen, & G. M. Burghardt (Eds.), *The cognitive animal: Empirical and theoretical perspectives on animal cognition* (pp. 197-204). Cambridge, MA: MIT Press.

Brannon, E. M., Wusthoff, C. J., Gallistel, C. R., & Gibbon, J. (2001). Numerical subtraction in the pigeon: Evidence for a linear subjective number scale. *Psychological Science, 12,* 238-243.

Breukelaar, J. W. C., & Dalrymple-Alford, J. C. (1998). Timing ability and numerical competence in rats. *Journal of Experimental Psychology: Animal Behavior Processes, 24,* 84-97.

Burns, R. A., Dunkman, J. A., Jr., & Detloff, S. L. (1999). Ordinal position in the serial learning of rats. *Animal Learning and Behavior, 27,* 272-279.

Burns, R. A., Goettl, M. E., & Burt, S. T. (1995). Numerical discrimination with arrhythmic serial presentations. *Psychological Record, 45,* 95-104.

Butterworth, B. (1999). *The mathematical brain*. London: Macmillan.

Call, J. (2000). Estimating and operating on discrete quantities in orangutans (*Pongo pygmaeus*). *Journal of Comparative Psychology, 114,* 136-147.

Cantlon, J. F., & Brannon, E. M. (2006). Shared system for ordering small and large numbers in monkeys and humans. *Psychological Science, 17,* 401-406.

Cantlon, J. F., & Brannon, E. M. (2007). How much does number matter to a monkey (*Macaca mulatta*)? *Journal of Experimental Psychology: Animal Behavior Processes, 33,* 32-41.

Capaldi, E. J., & Miller, D. J. (1988). Counting in rats: Its functional significance and the independent cognitive processes that constitute it. *Journal of Experimental Psychology: Animal Behavior Processes, 14,* 3-17.

Carey, S. (2004). Bootstrapping and the origin of concepts. *Daedalus, 133,* 59-68.

Church, R. M., & Meck, W. H. (1984). The numerical attribute of stimuli. In H. L. Roitblatt, T. G. Bever, & H. S. Terrace (Eds.), *Animal Cognition* (pp. 445-464). Hillsdale, NJ: Lawrence Erlbaum Associates.

D’Amato, M. R., & Colombo, M. (1988). Representation of serial order in monkeys (*Cebus apella*). *Journal of Experimental Psychology: Animal Behavior Processes, 14,* 131–139.

Davis, H. (1984). Discrimination of the number three by a raccoon (*Procyon lotor*). *Animal Learning and Behavior, 12,* 409-413.

Davis, H. (1993). Numerical competence in animals: Life beyond Clever Hans. In S. T. Boysen & E. J. Capaldi (Eds.), *The development of numerical competence: Animal and human models* (pp. 109-125). Hillsdale, NJ: Lawrence Erlbaum Associates.

Davis, H., & Albert, M. (1986). Numerical discrimination by rats using sequential auditory stimuli. *Animal Learning and Behavior, 14,* 57-59.

Davis, H., & Albert, M. (1987). Failure to transfer or train a numerical discrimination using sequential visual stimuli in rats. *Bulletin of the Psychonomic Society, 25,* 472-474.

Davis, H., & Bradford, S. A. (1986). Counting behavior by rats in a simulated natural environment. *Ethology, 73,* 265-280.

Davis, H., MacKenzie, K. A., & Morrison, S. (1989). Numerical discrimination by rats (*Rattus norvegicus*) using body and vibrissal touch. *Journal of Comparative Psychology, 103,* 45-53.

Davis, H., & Memmott, J. (1982). Counting behavior in animals: A critical evaluation. *Psychological Bulletin, 92,* 547-571.

Davis, H., & Pérusse, R. (1988). Numerical competence in animals: Definitional issues, current evidence, and a new research agenda. *Behavioral and Brain Sciences, 11,* 561-615.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition, 44,* 1-42.

Dehaene, S. (1997). *The number sense: How the mind creates mathematics.* New York: Oxford University Press.

Dehaene, S. (2003). The neural basis of the Weber-Fechner law: A logarithmic mental number line. *Trends in Cognitive Sciences, 7,* 145-147.

Dehaene, S., & Akhavein, R. (1995). Attention, automaticity, and levels of representation in number processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 314-326.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General, 122,* 371-396.

Dehaene, S., & Changeux, J. (1993). Development of elementary numerical abilities: A neuronal model. *Journal of Cognitive Neuroscience, 5,* 390-407.

Dehaene, S., Dehaene-Lambertz, G., & Cohen, L. (1998). Abstract representations of numbers in the animal and human brain. *Trends in Neurosciences, 21,* 355-361.

Dehaene, S., & Mehler, J. (1992). Cross-linguistic regularities in the frequency of number words. *Cognition, 43,* 1-29.

Dooley, G. B., & Gill, T. (1977). Acquisition and use of mathematical skills by a linguistic chimpanzee. In D. M. Rumbaugh (Ed.), *Language learning by a chimpanzee: The Lana project* (pp. 247-260). New York: Academic Press.

Emmerton, J., Lohmann, A., & Niemann, J. (1997). Pigeons' serial ordering of numerosity with visual arrays. *Animal Learning and Behavior, 25,* 234-244.

Evans, T. A., Beran, M. J., Harris, E. H., & Rice, D. (2008). *Quantification of food items by capuchin monkeys (*Cebus apella*): The influence of presentation style and the role of attention*. Manuscript submitted for publication.

Feigenson, L., & Carey, S. (2003). Tracking individuals via object-files: Evidence from infants' manual search. *Developmental Science, 6,* 568-584.

Feigenson, L., & Carey, S. (2005). On the limits of infants' quantification of small object arrays. *Cognition, 97,* 295-313.

Feigenson, L., Carey, S., & Hauser, M. (2002). The representations underlying infants' choice of more: Object files versus analog magnitudes. *Psychological Science, 13,* 150-156.

Feigenson, L., Carey, S., & Spelke, E. (2002). Infants' discrimination of number vs. continuous extent. *Cognitive Psychology, 44,* 33-66.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences, 8,* 307-314.

Fernandes, D. M., & Church, R. M. (1982). Discrimination of the number of sequential events by rats. *Animal Learning and Behavior, 10,* 171-176.

Ferster, C. B. (1964). Arithmetic behavior in chimpanzees. *Scientific American, 210,* 98-106.

Fetterman, J. G. (1993). Numerosity discrimination: Both time and number matter. *Journal of Experimental Psychology: Animal Behavior Processes, 19,* 149-164.

Fias, W., & Verguts, T. (2004). The mental number line: Exact and approximate. *Trends in Cognitive Science, 8,* 447-448.

Flombaum, J. I., Junge, J. A., & Hauser, M. D. (2005). Rhesus monkeys (*Macaca mulatta*) spontaneously compute addition operations over large numbers. *Cognition, 97,* 315-325.

Fuson, K. C., Pergament, G. G., & Lyons, B. G. (1985). Collection terms and preschoolers' use of the cardinality rule. *Cognitive Psychology, 17,* 315-323.

Gallistel, C. R. (1993). A conceptual framework for the study of numerical estimation and arithmetic reasoning in animals. In S. T. Boysen & E. J. Capaldi (Eds.), *The development of numerical competence: Animal and human models* (pp. 211-223). Hillsdale, NJ: Lawrence Erlbaum Associates.

Gallistel, C. R. (1999). Can a decay process explain the timing of conditioned responses? *Journal of the Experimental Analysis of Behavior, 71,* 264-271.

Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition, 44,* 43-74.

Gallistel, C. R., & Gelman, R. (2000). Non-verbal numerical cognition: From reals to integers. *Trends in Cognitive Sciences, 4,* 59-65.

Gast, H. (1957). The handling of numbers and number configurations in early childhood. *Zeitschrift für Psychologie, 161,* 1-90.

Gelman, R., & Gallistel, C. R. (1978). *The child's understanding of number*. Oxford, England: Harvard University Press.

Gelman, R., & Tucker, M. F. (1975). Further investigations of the young child's conception of number. *Child Development, 46,* 167-175.

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review, 84,* 279-325.

Gibbon, J. (1981). On the form and location of the psychometric bisection function for time. *Journal of Mathematical Psychology, 24,* 58-87.

Harris, E. H., & Washburn, D. A. (2008). *Subitizing and the iconic buffer*. Manuscript in preparation.

Hauser, M. D., & Carey, S. (2003). Spontaneous representations of small numbers of objects by rhesus macaques: Examinations of content and format. *Cognitive Psychology, 47,* 367-401.

Hauser, M. D., Carey, S., & Hauser, L. B. (2000). Spontaneous number representation in semi-free-ranging rhesus monkeys. *Proceedings of the Royal Society of London B, 267,* 829-833.

Hauser, M. D., Dehaene, S., Dehaene-Lambertz, G., & Patalano, A. L. (2002). Spontaneous number discrimination of multi-format auditory stimuli in cotton-top tamarins (*Saguinus oedipus*). *Cognition, 86,* B23-B32.

Hayes, J. J., & Nissen, C. H. (1971). Higher mental functions of a home-raised chimpanzee. In A. M. Schrier & F. Stollnitz (Eds.), *Behavior of non-human primates.* New York: Academic Press.

Hicks, L. H. (1956). An analysis of number-concept formation in the rhesus monkey. *Journal of*

Comparative and Physiological Psychology, 49, 212-218.

Honig, W. K., & Matheson, W. R. (1995). Discrimination of relative numerosity and stimulus mixture by pigeons with comparable tasks. *Journal of Experimental Psychology: Animal Behavior Processes, 21,* 348-363.

Honig, W. K., & Stewart, K. E. (1989). Discrimination of relative numerosity by pigeons. *Animal Learning and Behavior, 17,* 134-146.

Jensen, E. M., Reese, E. P., & Reese, T. W. (1950). The subitizing and counting of visually presented fields of dots. *Journal of Psychology: Interdisciplinary and Applied, 30,* 363-392.

Jordan, K. E., & Brannon, E. M. (2006). Weber's Law influences numerical representations in rhesus macaques (*Macaca mulatta*). *Animal Cognition, 9,* 159-172.

Jordan, K. E., & Brannon, E. M., & Gallistel, C. R. (2006). The multisensory representation of number in infancy. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 103,* 3486-3489.

Jordan, K. E., Brannon, E. M., Logothetis, N. K., & Ghazanfar, A. A. (2005). Monkeys match the number of voices they hear to the number of faces they see. *Current Biology, 15,* 1034-1038.

Judge, P. G., Evans, T. A., & Vyas, D. K. (2005). Ordinal representation of numeric quantities by brown capuchin monkeys (*Cebus apella*). *Journal of Experimental Psychology: Animal Behavior Processes, 31,* 79-94.

Keen, R., & Machado, A. (1999). How pigeons discriminate the relative frequency of events. *Journal of the Experimental Analysis of Behavior, 72,* 151-175.

Kilian, A., Yaman, S., von Fersen, L., & Güntürkün, O. (2003). A bottlenose dolphin discriminates visual stimuli differing in numerosity. *Learning and Behavior, 31,* 133-142.

Klahr, D. (1973). Quantification processes. In W. G. Chase (Ed.), *Visual information processing* (pp. 527-546). New York: Academic Press.

Klahr, D., & Wallace, J. G. (1973). The role of quantification operators in the development of conservation of quantity. *Cognitive Psychology, 4,* 301-327.

Koechlin, E., Dehaene, S., & Mehler, J. (1998). Numerical transformations in five-month-old human infants. *Mathematical Cognition, 3,* 89–104.

Koehler, O. (1950). The ability of birds to "count." *Bulletin of Animal Behaviour, 9,* 41-45.

Le Corre, M., & Carey, S. (2007). One, two, three, four, nothing more: An investigation of the conceptual sources of the verbal counting principles. *Cognition, 105,* 395-438.

Lewis, K. P., Jaffe, S., & Brannon, E. M. (2005). Analog number representations in mongoose lemurs (*Eulemur mongoz*): Evidence from a search task. *Animal Cognition, 8,* 247-252.

Lipton, J. S., & Spelke, E. S. (2003). Origins of number sense: Large-number discrimination in human infants. *Psychological Science, 14,* 396-401.

Logan, G. D., & Zbrodoff, N. J. (2003). Subitizing and similarity: Toward a pattern-matching theory of enumeration. *Psychonomic Bulletin & Review, 10,* 676-682.

Machado, A., & Rodrigues, P. (2007). The differentiation of response numerosities in the pigeon. *Journal of the Experimental Analysis of Behavior, 88,* 153-178.

Mandler, G., & Shebo, B. J. (1982). Subitizing: An analysis of its component processes. *Journal of Experimental Psychology: General, 111,* 1-22.

Matsuzawa, T. (1985). Use of numbers by a chimpanzee. *Nature, 315,* 57-59.

McCloskey, M., & Caramazza, A. (1987). Cognitive mechanisms in normal and impaired number processing. In G. Deloche & X. Seron (Eds.), *Mathematical disabilities: A* *cognitive neuropsychological perspective* (pp. 201-219). Hillsdale, NJ: Lawrence Erlbaum Associates.

McGonigle, B. (1988). Is it the thought that counts? *Behavioral and Brain Sciences, 11*, 593-594.

Mechner, F. (1958). Probability relations within response sequences under ratio reinforcement. *Journal of the Experimental Analysis of Behavior, 1,* 109-121.

Mechner, F., & Guevrekian (1962). Effects of deprivation upon counting and timing in rats. *Journal of the Experimental Analysis of Behavior, 5,* 463-466.

Meck, W. H., & Church, R. M. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behavior Processes, 9,* 320-324.

Menzel, E. W., Jr. (1960). Selection of food size in the chimpanzee and comparison with human judgments. *Science, 131,* 1527-1528.

Mitchell, R. W., Yao, P., Sherman, P. T., & O'Regan, M. (1985). Discriminative responding of a dolphin (*Tursiops truncatus*) to differentially rewarded stimuli. *Journal of Comparative Psychology, 99*, 218-225.

Morgan, C. L. (1914). *Introduction to comparative psychology* (revised new edition). New York: Scribner's.

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgments of numerical inequality. *Nature, 215,* 1519-1520.

Murofushi, K. (1997). Numerical matching behavior by a chimpanzee (*Pan troglodytes*): Subitizing and analogue magnitude estimation. *Japanese Psychological Research, 39,* 140-153.

Nevin, J. A. (1988). Reinforcement schedules and "numerical competence". *Behavioral and Brain Sciences, 11*, 594-595.

Nieder, A., Diester, I., & Tudusciuc, O. (2006). Temporal and spatial enumeration processes in the primate parietal cortex. *Science, 313,* 1431-1435.

Nieder, A., Freedman, D. J., & Miller, E. K. (2002). Large number discrimination in 6-month-old infants. *Science, 297,* 1708-1711.

Nieder, A., & Miller, E. K. (2004). Analog numerical representations in rhesus monkeys: Evidence for parallel processing. *Journal of Cognitive Neuroscience, 16,* 889-901.

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,* 325-339.

Olthof, A., & Roberts, W. A. (2000). Summation of symbols by pigeons (*Columba livia*): The importance of number and mass of reward items. *Journal of Comparative Psychology, 114,* 158-166.

Oyama, T., Kikuchi, T., & Ichihara, S. (1981). Span of attention, backward masking, and reaction time. *Perception & Psychophysics, 29,* 106-112.

Pepperberg, I. M. (1994). Numerical competence in an African grey parrot (*Psittacus erithacus*). *Journal of Comparative Psychology*, *108*, 36-44.

Pepperberg, I. M. (2006). Grey parrot (*Psittacus erithacus*) numerical abilities: Addition and further experiments on a zero-like concept. *Journal of Comparative Psychology, 120,* 1-11.

Pepperberg, I. M., & Gordon, J. D. (2005). Number comprehension by a Grey parrot (*Psittacus erithacus*), including a zero-like concept. *Journal of Comparative Psychology, 119,* 197-209.

Platt, J. R., & Johnson, D. M. (1971). Localization of position within a homogeneous behavior chain: Effects of error contingencies. *Learning and Motivation, 2,* 386-414.

Pylyshyn, Z. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition, 32,* 65-97.

Pylyshyn, Z. W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition, 80,* 127-158.

Roberts, W. A., & Mitchell, S. (1994). Can a pigeon simultaneously process temporal and numerical information? *Journal of Experimental Psychology: Animal Behavior Processes, 20,* 66-78.

Rugani, R., Regolin, L., & Vallortigara, G. (2007). Rudimental numerical competence in 5-day-old domestic chicks (*Gallus gallus*): Identification of ordinal position. *Journal of Experimental Psychology: Animal Behavior Processes, 33,* 21-31.

Rumbaugh, D. M., Hopkins, W. D., Washburn, D. A., & Savage-Rumbaugh, E. S. (1989). Lana chimpanzee learns to count by 'numath': A summary of a videotaped experimental report. *Psychological Record, 39,* 459-470.

Rumbaugh, D. M., Savage-Rumbaugh, S., & Hegel, M. T. (1987). Summation in the chimpanzee (*Pan troglodytes). Journal of Experimental Psychology: Animal Behavior Processes, 13,* 107-115.

Rumbaugh, D. M., Savage-Rumbaugh, E. S., & Pate, J. L. (1988). Addendum to 'Summation in the chimpanzee (*Pan troglodytes*)'. *Journal of Experimental Psychology: Animal Behavior Processes, 14,* 118-120.

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. 87-106). Hillsdale, NJ: Lawrence Erlbaum Associates.

Santos, L. R., Barnes, J. L., & Mahajan, N. (2005). Expectations about numerical events in four lemur species (*Eulemur fulvus, Eulemur mongoz, Lemur catta and Varecia rubra*). *Animal Cognition, 8*, 253-262.

Sawamura, H., Shima, K., & Tanji, J. (2002). Numerical representation for action in the parietal cortex of the monkey. *Nature, 415*, 918-922.

Seibt, U. (1982). Zahlbegriff und Zaehlverhalten bei Tieren: Neue Versuche und Deutungen [Number concepts and number behavior by animals: New experiments and interpretations]. *Zeitschrift für Tierpsychologie, 60,* 325-341.

Shumaker, R. W., Palkovich, A. M., Beck, B. B., Guagnano, G. A., & Morowitz, H. (2001). Spontaneous use of magnitude discrimination and ordination by the orangutan (*Pongo pygmaeus*). *Journal of Comparative Psychology, 115,* 385-391.

Silberberg, A., & Fujita, K. (1996). Pointing at smaller food amounts in an analogue of Boysen and Bernston’s (1995) procedure. *Journal of the Experimental Analysis of Behavior*, 66, 143-147.

Simon, T. J. (1997). Reconceptualizing the origins of number knowledge: A “non-numerical” account. *Cognitive Development, 12,* 349-372.

Simon, T. J., Hespos, S. J., & Rochat, P. (1995). Do infants understand simple arithmetic? A replication of Wynn (1992). *Cognitive Development, 10,* 253-269.

Simon, T. J., & Vaishnavi, S. (1996). Subitizing and counting depend on different attentional mechanisms: Evidence from visual enumeration in afterimages. *Perception & Psychophysics, 58*, 915-926.

Skinner, B. F. (1963). Operant behavior. *American Psychologist, 18,* 503-515.

Smirnova, A., Lazareva, O., & Zorina, Z. (2000). Use of number by crows: Investigation by matching and oddity learning. *Journal of the Experimental Analysis of Behavior, 73,* 163-176.

Smith, B. R., Piel, A. K., & Candland, D. K. (2003). Numerity of a socially housed hamadryas baboon (*Papio hamadryas*) and a socially housed squirrel monkey (*Saimiri sciureus*). *Journal of Comparative Psychology, 117,* 217-225.

Starkey, P., & Cooper, R. G. (1980). Perception of numbers by human infants. *Science, 210,* 1033-1035.

Starkey, P., Spelke, E. S., & Gelman, R. (1983). Detection of intermodal numerical correspondences by human infants. *Science, 222,* 179-181.

Starkey, P., Spelke, E. S., & Gelman, R. (1990). Numerical abstraction by human infants. *Cognition, 36,* 97-127.

Stocklin, P. L. (1999). 'Ordering of the numerosities 1 to 9 by monkeys': Comment and reply. *Science, 283,* 1851-1852.

Strauss, M. S., & Curtis, L. E. (1981). Infant perception of numerosity. *Child Development, 52,* 1146-1152.

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,* 73-85.

Terrace, H. S., Son, L. K., & Brannon, E. M. (2003). Serial expertise of rhesus macaques. *Psychological Science, 14,* 66–73.

Terrell, D. F., & Thomas, R. K. (1990). Number-related discrimination and summation by squirrel monkeys (*Saimiri sciureus sciureus* and *S. boliviensus boliviensus*) on the basis of the number of sides of polygons. *Journal of Comparative Psychology, 104,* 238-247.

Thomas, R. K. (1984). Are radical and cognitive behaviorism incompatible? *Behavioral and Brain Sciences, 7*, 650.

Thomas, R. K. (1988). To honor Davis & Pérusse and repeal their glossary of processes of numerical competence. *Behavioral and Brain Sciences, 11*, 600.

Thomas, R. K., Fowlkes, D., & Vickery, J. D. (1980). Conceptual numerousness judgments by squirrel monkeys. *American Journal of Psychology, 93*, 247-257.

Thomas, R. K., & Lorden, R. B. (1993). Numerical competence in animals: A conservative view. In S. T. Boysen & E. J. Capaldi (Eds.), *The development of numerical competence: Animal and human models* (pp. 127-147). Hillsdale, NJ: Lawrence Erlbaum Associates.

Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review, 101,* 80-102.

Uller, C., Carey, S., Huntley-Fenner, G., & Klatt, L. (1999). What representations might underlie infant numerical knowledge? *Cognitive Development, 14,* 1-36.

Uller, C., Hauser, M., & Carey, S. (2001). Spontaneous representation of number in cotton-top tamarins (*Saguinus oedipus*). *Journal of Comparative Psychology, 115,* 248-257.

Uller, C., Jaeger, R., Guidry, G., & Martin, C. (2003). Salamanders (*Plethodon cinereus*) go for more: Rudiments of number in an amphibian. *Animal Cognition, 6,* 105-112.

Uller, M. C. (1997). *Origins of numerical concepts: A comparative study of human infants and nonhuman primates*. Unpublished doctoral dissertation, Massachusetts Institute of

Technology, Cambridge.

Van Loosbroek, E., & Smitsman, A. W. (1990). Visual perception of numerosity in infancy. *Developmental Psychology, 26,* 916-922.

Warrington, E. K. (1982). The fractionation of arithmetical skills: A single case study. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 34A,* 31-51.

Washburn, D. A. (1994). Stroop-like effects for monkeys and humans: Processing speed or strength of association? *Psychological Science, 5,* 375–379.

Washburn, D. A., & Rumbaugh, D. M. (1991). Ordinal judgments of numerical symbols by macaques (*Macaca mulatta*). *Psychological Science, 2,* 190–193.

Wender, K. F., & Rothkegel, R. (2000). Subitizing and its subprocesses. *Psychological Research, 64,* 81-92.

West, R. E., & Young, R. J. (2002). Do domestic dogs show any evidence of being able to count? *Animal Cognition, 5,* 183-186.

Whalen, J., Gallistel, C. R., & Gelman, R. (1999). Nonverbal counting in humans: The psychophysics of number representation. *Psychological Science, 10,* 130-137.

Wolters, G., van Kempen, H., Wijlhuizen, G. (1987). Quantification of small numbers of dots: Subitizing or pattern recognition? *American Journal of Psychology, 100,* 225-237.

Wynn, K. (1990). Children's understanding of counting. *Cognition, 36,* 155-193.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature, 358,* 749-750.

Wynn, K. (1996). Infants' individuation and enumeration of actions. *Psychological Science, 7,* 164-169.

Xia, L., Emmerton, J., Siemann, M., & Delius, J. D. (2001). Pigeons (*Columba livia*) learn to link numerosities with symbols. *Journal of Comparative Psychology*, *115*, 83-91.

Xia, L., Siemann, M., & Delius, J. D. (2000). Matching of numerical symbols with number of responses by pigeons. *Animal Cognition, 3,* 35-43.

Xu, F. (2003). Numerosity discrimination in infants: Evidence for two systems of representations. *Cognition, 89,* B15-B25.

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition, 74*, B1-B11.

Xu, F., Spelke, E. S., & Goddard, S. (2005). Number sense in human infants. *Developmental Science, 8*, 88-101.

Chapter 2: What do Arabic numerals mean to macaques?[1]

Abstract

In the past, rhesus macaques have demonstrated an ability to use Arabic numerals to facilitate performance in a variety of tasks. However, it remained unclear whether they understood the absolute as well as the relative values of numerals. In Experiment 1, numeral-trained macaques picked the largest stimuli when presented with pair-wise comparisons involving numerals and analog quantities. In Experiment 2, macaques were provided with numeral cues indicating the number of times a behavior could be performed in one location for a reward. Three of the four monkeys performed above chance, but they often erred by performing more behaviors than indicated. The results of these studies indicate that the monkeys have knowledge of the approximate quantities represented by each numeral.

[1] This chapter has been submitted for publication as: Harris, E. H., Gulledge, J. P., Beran, M. J., & Washburn, D. A. (2008). What do Arabic numerals mean to macaques?

A range of nonhuman animals including dolphins (Mitchell, Tao, Sherman, & O'Regan, 1985), pigeons (Olthof & Roberts, 2000; Xia, Emmerton, Siemann, & Delius, 2001), a parrot (Pepperberg, 1994, 2006), squirrel monkeys (Olthof, Iden, & Roberts, 1997), capuchin monkeys (Beran et al., 2008), rhesus monkeys (Beran, Beran, Harris, & Washburn, 2005; Washburn & Rumbaugh, 1991) and chimpanzees (Beran & Rumbaugh, 2001; Biro & Matsuzawa, 2001; Boysen & Berntson, 1989, 1995; Matsuzawa, 1985; Murofushi, 1997; Rumbaugh, Hopkins, Washburn, & Savage-Rumbaugh, 1989) have learned to make numerical judgments using arbitrary symbols that represent quantities. The advantage to using arbitrary symbols, such as Arabic numerals, in numerical tasks rather than food items or analog stimuli is that the symbols provide no inherent non-numerical cues such as surface area, density, or complexity to indicate the relation between one quantity and another.

Over the last decade, the rhesus monkeys in our laboratory have participated in a variety of studies aimed at assessing their ability to perform numerical tasks using Arabic numerals. This research began when Washburn and Rumbaugh (1991) presented rhesus monkeys with pairs of the numerals 0 through 9 and reinforced them with a corresponding number of pellets for choosing either of the numerals. The monkeys learned to choose the larger numeral, and they performed accurately even when presented with probe trials of unfamiliar pairings of numerals. None of the probe pairings could be solved on the basis of logical transitivity. For instance, knowledge that $8 > 7$ and $8 > 6$ does not provide sufficient information to conclude that $7 > 6$. When the animals were later presented with arrays of up to five numerals, they tended to select stimuli in the correct reverse ordinal sequence.

Across the years, dozens of additional monkeys were trained and tested at our laboratory or at the Ames Research Center (Moffett Field, CA) using our software and the protocols originally described for the two animals by Washburn and Rumbaugh (1991). For these animals, the relative numerousness judgment task with Arabic numerals, described above, was administered as one of many training tasks to prepare the animals for subsequent cognitive studies. We have never reported the data from these training sessions, but the findings replicated the earlier results. We currently have summary data for 66 of those monkeys. After receiving training with randomly paired numerals (0 to 9, as described above and by Washburn & Rumbaugh, 1991), with proportional rewards for whichever numeral they selected, these animals averaged 84% accuracy on familiar (trained) pairings. On the first presentation of novel test trials—pairings of numerals the animals had never seen before—the monkeys were correct on a total of 367 of the 462 probe trials (7 novel probes for each of the 66 animals). This accuracy level (79%) is substantially and significantly in excess of what would be predicted by chance, and is in fact very near the monkeys' levels of accuracy for familiar, over-trained pairings.

Despite these demonstrations of numerical ability, the following question remained: What exactly do these animals know about Arabic numerals? In the Washburn and Rumbaugh (1991) study, for example, it is possible that the monkeys learned a complex matrix involving the relative difference and degree of difference between all possible pairs of numerals. For example, knowledge that 8 is greater than 7 by one pellet (the smallest unit of difference) and greater than 6 by two pellets (the second smallest unit of difference) would allow the monkeys to solve a novel pairing of 8 and 7 and also 7 and 6. A different explanation is that the monkeys gained an understanding of the absolute quantity of pellets represented by each number. Perhaps the most parsimonious explanation is that the animals associated each numeral with a different hedonic value based on how much food was presented for that numeral. In others words, the largest numerals evoked the strongest positive hedonic states.

There is some evidence that monkeys may understand the quantitative values represented by Arabic numerals. In a study of Stroop-like effects (Washburn, 1994), rhesus monkeys at our laboratory learned to select the larger of two arrays of 1 to 9 letters (e.g., to select five As rather than four Cs). When the arrays of letters were replaced with arrays of numerals, incongruous numerals (e.g., four 1s versus two 5s) disrupted performance and congruous numerals (four 5s versus two 4s) did not. In other words, it was more difficult for the monkeys to choose the array with the most stimuli when that array was composed of the smaller numeral than when it was composed of the larger numeral. This effect suggests that these monkeys processed the quantitative meanings of the numerical symbols automatically because of their prior training with these numerals, despite the fact that these meanings were irrelevant to the task.

Experiment 1 of the current study further tests the hypothesis that number-trained monkeys understand the absolute as well as the relative values of numerals by presenting them with pair-wise comparisons involving numerals and analog quantities. If the monkeys had originally learned a complex matrix of values using knowledge of the relative difference and degree of relative difference between pairs of numerals, then they should be incapable of comparing symbols with actual quantities. Conversely, if the monkeys acquired knowledge of the absolute quantity of pellets represented by each Arabic numeral, then they might be able to compare symbols with analog dot arrays.

Additionally, the study includes a test of whether it is number or hedonic value that determines the monkeys' behavior. Brannon and Terrace (1998) criticized the proportional reinforcement procedure previously used in some numerical studies (e.g., Washburn & Rumbaugh, 1991) because it confounds numerousness and hedonic value, making it possible that the monkeys were not responding to the stimuli based upon numerosity, but rather on the richer reinforcement history provided by certain numerals that led to larger numbers of food pellets. That is, Brannon and Terrace (1998) suggested that the monkeys in our laboratory chose 7 instead of 6, not because 7 is "more" than 6, but because 7 is "better" than 6. We acknowledge this possibility, but note that monkeys could also perceive that 7 is better than 6 because 7 is more likely to be reinforced, as in studies like Brannon and Terrace (1998) that used probabilistic (rather than proportional) reinforcement in which the animal is rewarded only for selecting the correct number. To address this criticism empirically, five monkeys in Experiment 1 received proportional rewards for every selection and a sixth monkey (Hank) was rewarded with a proportional number of pellets for numerals (e.g., picking the 4 netted four pellets) but probabilistic rewards for dot selections (e.g., correctly picking the bigger array always netted one pellet).

Experiment 2 was designed to further assess the monkeys' use of absolute numerical knowledge using a sequential task in which the monkeys were required to enumerate their own responses. In a previous study we used a similar method to investigate the ability of four of our number-trained rhesus monkeys to use Arabic numeral cues to discriminate between different series of maze trials and anticipate the final trial in each series (Harris & Washburn, 2005). The monkeys were trained on a computerized task consisting of three reinforced maze trials followed by one nonreinforced trial. The goal of the maze was an Arabic numeral 3, which corresponded to the number of reinforced maze trials in the series. Two of the four monkeys developed a "slow, fast, faster, slow" pattern, which suggested they were anticipating the final nonreinforced trial. The other two monkeys performed gradually slower on each trial in a series, which made it impossible to ascertain whether or not they were predicting precisely when the final trial would occur.

During testing, two monkeys were given probe series of the numerals 2 and 4 intermixed with the familiar numeral 3 series and the remaining two monkeys were given probe series of the numerals 2 through 8. As was true during training, the Arabic numeral displayed in the maze corresponded to the number of reinforced trials that would occur before one nonreinforced trial. The monkeys receiving probe trials of the numerals 2 and 4 showed some generalization to the new numerals and developed a pattern of performing more slowly on the nonreinforced trials than the reinforced trials, indicating the use of the changing target numeral cues to anticipate those final nonreinforced trials. The monkeys receiving probe trials of the numerals 2 through 8 did not use the changing numeral to predict precisely when the nonreinforced trial would occur in each series, but they did incorporate the changing numerals into their strategy by performing faster overall on series with greater target numerals.

The Harris and Washburn (2005) study provided evidence that number-trained rhesus monkeys could use Arabic numerals as a cue to facilitate performance on a task involving sequential responses, also known as a “constructive” enumeration task (Beran & Rumbaugh, 2001; Xia, Siemann, & Delius, 2000). However, the pattern established by two of the monkeys during training of performing gradually slower on each trial in a series, and the failure of the monkeys receiving probe trials of the numerals 2 through 8 to generalize the pattern learned during training to new target numerals highlighted the need for a task that specifically addressed the monkeys’ understanding of when a series is finished. Thus, in Experiment 2, monkeys were provided with Arabic numeral cues indicating the number of times a behavior could be performed in one location for a reward. After receiving all of the possible rewards from one location the monkeys could behaviorally indicate that the series was complete by moving on to a second location. This design allowed us to assess their understanding of the cardinal value of the numerals. Together, these two experiments provided us with a greater understanding of what Arabic numerals mean to the macaques in our laboratory.

Experiment 1: Do monkeys know, for example, that 4 > ● ● ● ?

Method

Subjects. Six male rhesus monkeys (*Macaca mulatta*) were tested in this study. The monkeys (age range 4 to 16 years) had previously been trained following the procedures described elsewhere (Rumbaugh, Richardson, Washburn, Savage-Rumbaugh, & Hopkins, 1989) to manipulate a joystick so as to control a computer-graphic cursor in response to stimuli displayed on a computer screen. The animals were not deprived of food or water and had continuous access to the apparatus and computerized tasks so that they could work or rest ad libitum. Each of the monkeys had been tested in numerous experiments prior to the present study, and each had previously learned to respond to Arabic numerals in accordance with the number of pellets associated with each (see Washburn & Rumbaugh, 1991, for the details of this task and the procedure by which all of these monkeys learned to select the greater of any pair or array of Arabic numerals). That is, each monkey could generally select the larger of any two Arabic numerals (0 to 9) to receive the corresponding number of pellets, and could generally select the array of letters or numerals with the most items (Washburn, 1994). Importantly, none of the monkeys had received any training to link directly these two dimensions of numerousness (e.g., no training to label four items with the numeral 4, or to pick three stimuli when presented with the numeral 3, or to determine whether the numeral 2 is greater than or less than some number of dots).

Apparatus. An analog joystick was connected to a computer that displayed stimuli on a 13-inch color monitor, presented auditory feedback through an external speaker/amplifier, and delivered 97-mg fruit-flavored chow pellets (Noyes, Lancaster, NH) via a pellet dispenser (Gerbrands 5210) and relay interface (ERA01 and PIO12, Keithley). The mounting and protection of this apparatus has been described in detail elsewhere (Rumbaugh et al., 1989; Washburn & Rumbaugh, 1992). Each monkey worked at a dedicated computerized test system, and the monkey could reach through the mesh of his home cage to manipulate the joystick and to retrieve pellets.

Task. Each trial began with the cursor (a white "+") randomly positioned on the screen and a small (1.25 cm diameter) circle presented midscreen. Trials were initiated by manipulating the joystick so as to direct the cursor into the circle, whereupon the numerical stimuli were presented on either side of the cursor. Three stimulus conditions were used in these experiments. Some trials were numeral:numeral trials, in which two different Arabic numerals (1 to 9) were selected at random and positioned randomly to the left and right of the cursor. Other trials were dot:dot trials, in which two different arrays of 1 to 9 randomly positioned 2.5 cm diameter white dots were displayed, one array on each side of the cursor. During training, each trial was randomly determined to be a numeral:numeral or a dot:dot trial. During subsequent probe testing, some trials were numeral:dot trials in which a randomly selected Arabic numeral was presented on the screen with a randomly selected (but different) quantity of dots. The position of the stimuli was randomized for all conditions (i.e., the numeral did not always appear on the left for numeral:dot trials). The monkeys' choices and response times were recorded for each trial.

Training procedure, proportional reinforcement. Five of the monkeys (Murph, Lou, Baker, Gale, and Willie) were trained to criterion with a version of the task that delivered a number of reinforcements proportional to the numeral or analog dot array that was selected.

That is, an animal received seven pellets for picking the numeral 7 or an array of seven dots, received three pellets for picking the numeral 3 or an array of three dots, and so forth.

Training procedure, conditional reinforcement. To assess the hedonic criticism outlined above, we decided to implement two control conditions in the training and testing of Hank. First, Hank was reinforced proportionally for the numeral:numeral trials (as had been done throughout his prior test history), but he was only reinforced with a single pellet for correct dot:dot responses, irrespective of the number of dots in the array. Second, we withheld some of Hank's dot:dot training trials, so that he never received an array in which five dots was the smaller quantity. That is, every one of Hank's choices of five dots was correct and reinforced during training.

Probe test procedure. After the monkeys reached a criterion of at least 76% accuracy on numeral:numeral trials and on dot:dot trials, the 72 possible novel numeral:dot trials were interspersed randomly within the next session. Note that one could claim that more than 72 novel numeral:dot trials exist, given that the position of the dots and the numeral were randomized each trial; however, the first exposure of each numeral with each quantity of dots (irrespective of position) was considered a probe trial for this study. These trials were reinforced in the same way described above for the training conditions: proportional reinforcement for numerals, probabilistic reinforcement for Hank's dot selections, and all selections of the smaller numeral or quantity of dots resulting only in a 1-second buzz and no food reward.

Results

Each monkey achieved the training criterion in fewer than 1,500 trials (about one day of testing). Overall accuracy averaged 84% for numeral:numeral comparisons and 82% for dot:dot comparisons. As was reported by Washburn (1994) and Brannon and Terrace (1998, 2000), accuracy and response time varied as a function of the absolute difference between the quantities or numerals (e.g., 7:2 responses were faster and more accurate than 3:2 responses or three dot:two dot responses).

In the probe-test phase of the study, performance for each of the six monkeys was significantly better than chance ($p < .05$, binomial test), with the novel numeral:dot trials averaging 81.5% accuracy (see Table 2.1). No reliable differences were observed in the probe-trial tests between the numeral:numeral, dot:dot, and numeral:dot conditions.

Particular attention should be directed to Hank's responses in this probe-test phase. In the initial 36 novel numeral:dot trials, Hank showed a reliable numeral bias, and thus was correct on 100% of the trials in which the numeral was larger but only 6% of the trials in which the dot array was larger. Following one additional day of testing on numeral:numeral, dot:dot, and these 36 numeral:dot trials, he was tested on the final 36 novel probes. Hank was correct on 89% of these novel probes, and accuracy on the other trial types remained high (93%). It is also noteworthy that Hank was correct on 7 of the 8 (88%) novel trials in which an array of five dots was paired with a larger array or numeral.

For all of the monkeys, most of the errors on the numeral-dot probe trials were made on trials in which the numerical distance (i.e., the difference between the numerical values of the numeral and dot array) of the two stimuli was very small. Analysis of variance revealed a significant effect of numerical distance for the accuracy data, $F\ (7, 28) = 10.02$, $p < .05$. A post-hoc analysis utilizing a Tukey-test for Honestly Significant Difference (HSD) revealed that performance for the numerical distance of 1 was significantly different from performance for numerical distances of 2 and larger, performance for the distance of 2 was significantly different from performance for the distances of 4, 5, and 8, and performance for the distance of 3 was significantly different from performance for the distances of 4 and 8.

Discussion

The monkeys accurately interdigitated Arabic numerals and random arrays of dots, even on the first exposure to these trials. This indicates that the monkeys were not relying solely on a complex matrix of two-choice discriminations learned during the training phase. Knowledge that the numeral 7 is the correct choice when presented with the numeral 6 and that an array of seven dots is the correct choice when presented with six dots is not sufficient information to compare the numeral 7 and six dots.

Additionally, the data from Hank indicate that this ability is not based solely on the hedonic value of the numerals. Although Hank initially favored all numerals, probably because of their substantial advantage in terms of reinforcement history, he did come to respond at levels significantly better than chance to first-trial presentations of stimulus pairs in which numerical value opposed hedonic value. For example, Hank, like the other monkeys, responded that four dots is greater than the numeral 3, even though the numeral 3 was associated with a rich reinforcement history whereas arrays of four dots were only occasionally reinforced, and then with only a single pellet. This point is further supported by the observation that Hank did not base his responses solely on “probability of reward” either. Although arrays of five dots were always reinforced during training, Hank correctly selected larger arrays or Arabic numerals greater than 5 on 88% of the subsequent probe trials.

Hank’s data are also interesting because they indicate that he was able to use the ordinal information inherent in the dot quantities to perform the numeral and dot comparisons. Dot quantities are different from symbolic stimuli such as numerals because they have visible properties that relate to their ordinal position in a sequence (i.e., the quantity three consists of fewer dots than the quantity four). During training Hank was only reinforced with a single pellet for correct dot:dot responses and he never received an array in which five dots was the smaller quantity. Thus, he did not have an opportunity to associate specific quantities of pellets with specific dot quantities or to learn the order of the dot quantities 5 through 9 using reinforcement history. The fact that Hank was able to solve the novel trials in which an array of five dots was paired with a larger array or numeral indicates that he was responding to the inherent ordinal value present in the analog dot displays. These data match those from another study with rhesus monkeys in which those animals also spontaneously responded to analog dot displays in a comparison task on the basis of their ordinal relations (Harris, Beran, & Washburn, 2007).

Based on the combined data from all of the monkeys we are able to rule out a matrix of learned values and also hedonic value as the basis for responding to the novel numeral:dot comparisons. The data therefore suggest that the monkeys had acquired knowledge about the absolute quantity of things represented by each Arabic numeral and could, even on probe trials, compare accurately this represented quantity to a visible array of dots.

Another possibility that must be noted, however, is that performance reflected integration of two learned sequences instead of comparisons of quantity. Research indicates that monkeys trained to order two lists of four arbitrary stimuli (e.g., $A_1B_1C_1D_1$ and $A_2B_2C_2D_2$) immediately respond correctly at a greater than chance level when presented with comparisons of two items from different lists (e.g., A_1-C_2 or B_1-D_2; D’Amato & Colombo, 1988; Terrace, Son, & Brannon, 2003). It is possible that the monkeys in the current study perceived the numerals as one arbitrary list of stimuli and the dot quantities as another arbitrary list and ordered pairs of numerals and dots using only knowledge of their ordinal position. It is unlikely, however, that the differentially reinforced monkeys completely disregarded quantity information during the novel probe comparisons given that quantity information was readily available during training. We know that Hank did not represent the dot arrays as arbitrary stimuli in a list with no inherent order because he was able to solve novel numeral:dot probe trials, despite the fact that some dot pairs were withheld during training so he had no opportunity to learn the complete order of dot quantities by trial-and-error.

In fact, the use of pair-wise comparisons leaves open the question of whether or not any of the monkeys learned a complete ordered list of numerals or dot quantities. Knowing that 8 > 7 and 7 > 6 is not the same as knowing that 8 > 7 > 6. Subsequent studies have produced mixed findings pertaining to the formation of an ordered list based on pair-wise comparison training. When two of the monkeys in the current study were subsequently trained to order lists of Arabic numerals and arbitrary colors they showed no advantage with the numerals, which suggests that they had no previous representation of Arabic numerals as an ordered list (Harris et al., 2007). In another study, however, capuchin and rhesus monkeys from our laboratory were presented with random pairings of the Arabic numerals 0 through 9 and learned to choose the larger numeral when rewarded with one pellet for each correct choice. The monkeys were subsequently presented with arrays of 5 familiar numerals and arrays of 5 novel letters and both species performed better with the numerals. This indicates that they had learned a sequence of numerals during the pairwise comparison training despite the lack of quantity information (Beran et al., 2008). Regardless of the exact nature of the knowledge used in the current task, the monkeys were able to compare novel numerals and dot quantities based only on information acquired during randomly presented pairwise comparisons.

Overall, the data from this experiment provide insight regarding what the monkeys know about the numeric symbols with which they have experience. It appears that the numeral 4 does not simply mean “greater than 3 and less than 5” or “better than 3 and not as good as 5,” but rather it represents a quantity that can be compared directly and accurately to visible arrays of analog stimuli. However, it is important to note that performance suffered when the numerical distance between the numeral and dot quantity was small, which suggests that any quantity information was approximate rather than exact.

Experiment 2: Do monkeys know, for example, that 4 means four actions?

The monkeys in the sequential study by Harris and Washburn (2005) that was discussed in the introduction were clearly using the numeral values to alter their behavior on a sequence of maze trials. They solved nonreinforced trials more slowly than reinforced trials. They could have produced this effect in the way suggested by the authors, by solving the maze slowly in anticipation of a lack of reward when the number of reinforced trials performed matched the value of the target numeral. However, they could also have accomplished this by performing more slowly on each successive trial (without keeping track of the number of reinforced trials or even knowing the cardinal value of the target number) and resetting back to rapid responding after the nonreinforced trial. Indeed, two monkeys appeared to do this. It is important to note that even if this was the strategy, the monkeys were still using the numeral values to adjust performance speeds differentially, so that the slope of successive slowing was steeper when the target number was 3 than when it was 5.

An alternative procedure is required to allow a monkey to solve a maze N times and then behaviorally indicate, “I’m done.” We reasoned that by placing two target numerals in the maze, the monkeys could travel to the larger number the corresponding number of times and then indicate that he knew he had exhausted the pellets at this location by traveling to the smaller number. In contrast, if the monkeys know only ordinal and not absolute values corresponding to the numerals, then the animals have no basis for knowing when to stop responding to the larger of two numerals and to move instead to the smaller stimulus.

Method

Subjects. Four rhesus monkeys were available to be tested in this experiment. The animals (Hank, Gale, Willie, and Murph) had been trained previously to select between visual arrays or Arabic numerals and had participated in some of the experiments discussed above. All were familiar with moving a cursor through a two-dimensional maze (Harris & Washburn, 2005; Harris, Washburn, Beran, & Sevcik, 2007), although none had seen the task with two Arabic-numeral targets prior to this study.

Task. Each trial began with a white plus-sign cursor ("+", measuring 1.25 cm × 1.25 cm) presented midscreen against a black background. White rectangles were displayed on the screen to form a basic two-dimensional H-maze (see Figure 2.1). Two randomly selected Arabic numerals were presented in two terminus points of the maze, equidistant from the cursor. For each trial within a problem, these numeral positions remained constant. Each numeral was "baited" with the corresponding number of pellets. Moving the cursor into contact with either numeral resulted in the delivery of a pellet, unless the monkey had already earned the corresponding number of pellets for that problem (e.g., a 4 would only be reinforced four times in a problem). Additional responses to a numeral were scored as errors. When the monkeys made an error they received a negative buzzing sound and the cursor reset to the center of the screen in preparation for the next trial. The monkey could make as many errors as needed to obtain all the pellets for each problem. The problem ended automatically when all of the pellets that could be obtained had been earned. A new problem began immediately after the previous problem ended. New numerals and random positions for the numerals were generated for each new problem.

Procedure. Gale and Hank were trained for 200 problems in which the target numerals were always 2 and 3. Between problems, the location of the numerals was changed randomly. For each of these 200 problems, an ideal solution was to move the cursor through the maze to the 3 on three (and only three) trials and to the 2 on two (and only two) trials. Note that nothing constrained the animals to select the numerals in this order (i.e., touching the 2 twice and then the 3 thrice would also have been errorless performance, as would other combinations of responses that did not involve moving to a numeral more times than its value). After these 200 problems, another numeral (1 to 6) was introduced every 50 problems. For example, the Arabic numerals on Gale's problems 201-250 were 2, 3, or 5 and for Hank they were 2, 3, and 4.

An identical procedure was used for Willie and Murph, except that their first 200 training problems used 2 and 4 as targets. As above, an additional numeral was selected at random every 50 problems to be included in the stimulus pool.

Results

All four monkeys learned to complete problems during the initial training period, and three of the four animals generalized to new numerals when they were added to the sequence. Figure 2.2 shows the percentage of problems completed without error (i.e., without visiting a numeral more times than its value), relative to chance. Chance was computed separately for each monkey because each monkey received different number pairings. Gale, Willie, and Murph performed significantly better than chance across target-numeral pairings ($p < .05$, binomial test).

Hank showed a different pattern. During the initial training problems, he developed the strategy of alternating between the target numerals, starting with the larger numeral. That is, he learned that he could touch the 3, 2, 3, 2, 3, in sequence to end each trial without error. Of course, this was a perfectly acceptable strategy, but one that would not work when most other combinations of numerals were used as targets. Consequently, Hank's performance was statistically at chance levels on the test trials.

Examining the data for the other three animals, trials without error were seen at levels in excess of chance across target numerals and target-number ratios. Figure 2.3 depicts average performance across target numerals and Figure 2.4 depicts average performance across target-number ratios. Note that performance was essentially stable across ratios—consistently above and showing a different distribution than chance levels (as determined by Monte Carlo simulation). Performance was significantly better than chance ($p < .01$) at every ratio except 0.67 and 0.83 (each $p > .10$). Observed behavior was better simulated by an algorithm that selected numerals in proportion to their relative magnitudes (i.e., to be twice as likely to select 4 rather than 2 when they were paired together, versus having a .50 chance of selecting each numeral). However, even this simulation failed to capture the level of errorless trials that was observed with target:target ratios of 0.6 and greater. The monkeys performed significantly better than the relative amount simulation for the target pairings 5:3, 4:3, and 5:4 ($p < .05$). The monkeys' performance on these problems required knowledge beyond the relative magnitudes of the numerals.

To determine how the monkeys solved these problems we examined the pattern of responding and found several patterns that were routinely used by the monkeys in this task. Some of the problems were solved with a pattern we labeled as the "numeral pattern." This pattern involved clearing out the bigger numeral first before moving to the smaller numeral. For example, if the problem contained a numeral 2 and a numeral 3 the monkeys would contact the 3 on three consecutive trials before contacting the 2 on the last two trials.

Other problems were solved with a pattern we labeled as the "pellet pattern." This pattern involved contacting the numeral with the greatest number of remaining pellets on every trial in a problem. For example, when presented with a 2 and a 4, the monkey might touch 4, 4, 4, 2, 2, 4, in that order. After contacting the 4 on three trials the numeral 2 would have the most remaining pellets (one pellet for the numeral 4 and two for the numeral 2) so the monkeys might switch to the numeral 2. After contacting the numeral 2 twice, the numeral 4 would have the most remaining pellets (one pellet for the numeral 4 and zero for the numeral 2) so the monkeys might switch back to the 4 to finish the problem. It must be noted, however, that the same pattern for the numerals 2 and 4 could be obtained using a slightly different rationale. The monkeys could attempt to clear out the numeral 4 first, but move prematurely to the 2. If they knew they had exhausted the pellets at the 2 and the problem did not end, they could then move back to the 4 and retrieve the last pellet. This would produce the same pattern as choosing the numeral with the largest number of pellets on every trial. This rationale, however, would not produce the same pattern as the "pellet pattern" for other pairings such as 4 and 3.

The "alternating pattern" was scored as a special case of the pellet pattern. In the alternating pattern, the monkeys started with the larger numeral and alternated between the two numerals on each response until the pellets had been exhausted. So if they were presented with a 3 and a 2 they would touch 3, 2, 3, 2, 3, in that order. Note that this response pattern would never produce an errorless trial when the target numerals differed by more than one (e.g., 5 and

3). Problems in which the smaller number was cleared first (e.g., 1, 2, 2) or there was no predictable pattern were labeled as the “other pattern.”

We determined the proportion of errorless trials as a function of response pattern for Gale, Willie, and Murph. The proportion was approximately 5% for the numeral pattern, 20% for the pellet pattern, 20% for the alternating pattern, and 55% for the other pattern. We then computed how often the monkeys’ behavior would conform to these patterns by chance, given the numeral pairings that were used in the study. We used a computer simulation to perform this calculation.

Simulations were conducted by creating a computer program that responded to the same kinds of problems the monkeys saw. That is, we simulated the choices between numerals, not the maze-running itself. The simulation chose randomly between the two numerals available in the problem. Responding continued in this way until the trial was completed (i.e., until each numeral had been selected the corresponding number of times). Errors were calculated for the computer in the same way they were operationalized for the monkeys (e.g., choosing the 4 more than four times in a problem). The computer was tested with blocks of problems, as was done with the monkeys, but for purposes of generating the normal distribution, at least 10,000 blocks of trials were simulated for each possible pairing of numerals (1 to 6). One million trials were simulated in total.

After each block of problems, the proportion of trials completed without error was calculated, producing a sampling distribution of errorless trials that could be expected by chance alone. Each simulated errorless trial was also scored according to whether it matched the pellet pattern described above. In this way, we obtained statistical estimates of the likelihood by chance alone of selecting the numeral associated with the larger number of pellets on every trial (response) of a problem. Of course, the probability of this pattern of responses varied as a function of the numeral pairing. For 2:1 problems, 2/3 of the problems fit this pattern (i.e., only by selecting the 1 on the first response could one complete a trial without error and without following the pellet pattern). By comparison, only 7% of the errorless 6:5 trials fit this pattern by chance alone. Overall, only 20% of the trials that were completed without error by the computer simulation matched the pellet pattern.

In the monkeys' responses, 40% of the errorless trials involved touching the numeral on each trial that had the same or greater number of pellets remaining, including those trials in which the animals alternated between the numerals and thereby selected the numeral with the larger number of remaining pellets (i.e., the pellet pattern plus its variation, the alternating pattern). The 40% of errorless trials that actually fell into those two related categories significantly exceeded the chance level determined by the computer simulation ($p < .05$). This suggests that the monkeys were purposefully using this pattern to facilitate performance, and not just behaving at random with some of their behavior conforming to the pellet and alternating strategies by chance.

No other strategy was observed at levels in excess of chance. Analysis revealed that 8% of the simulated trials fit the "numeral pattern" (e.g., 5,5,5,5,5,2,2), which is a number statistically identical to the 5% the monkeys produced ($p > .10$). Errorless responses that fit the "alternating pattern" (e.g., 2,1,2) alone were even less probable by chance (3%), but the monkeys did not produce this subset of the "pellet pattern" at levels significantly in excess of chance ($p = .06$). Recall that the "alternating" variant of the pellet pattern could only produce errorless performance on five numeral pairings (1:2, 2:3, 3:4, 4:5, and 5:6).

Discussion

Three of the four animals performed better than chance, and better than would be expected if they knew only the relative magnitudes of the numerals. The pattern of responding indicated that the animals learned strategies to simplify the task, such as clear the larger numeral first and choose the numeral with the greatest amount of remaining pellets. Correct execution of these strategies required knowledge beyond ordinality (which numeral is bigger) or even ratio (the relative magnitude of the difference in proportions). For instance, the pellet strategy required the monkeys to know the absolute number of pellets remaining for each numeral.

Despite these performance strategies, errorless problems were still the minority. On most trials, the monkeys touched a target more times than was represented by the numeral. Of course, the memory demands of this task were substantial, requiring a monkey not only to keep track of how many times he had touched a specific target, but potentially also to remember how many times he had touched the other target, and in any case to reset these representations for each new problem. Under these demands it seems unreasonable to expect errorless performance on the vast majority of trials; still, the present data do not compel a conclusion that the monkeys were enumerating responses toward some exact and absolute quantity (e.g., move to the 3 exactly three times). Instead, the results indicate that the monkeys had an understanding of the approximate values represented by the numerals.

General Discussion

Over the past two decades, researchers have provided clear evidence that nonhuman primates can use Arabic numerals to perform a variety of tasks (e.g., Beran & Rumbaugh, 2001; Biro & Matsuzawa, 2001; Boysen & Berntson, 1989; 1995; Matsuzawa, 1985; Murofushi, 1997; Olthof et al., 1997; Rumbaugh et al., 1989; Washburn & Rumbaugh, 1991). Despite these impressive displays of numerical competence, it is difficult to ascertain exactly what numerals represent to these animals. When adult humans look at the Arabic numeral 4 they understand several things about that numeral. They understand that the numeral 4 is larger than the numeral 3 and smaller than the numeral 5 and that it symbolizes the quantity four. They also understand that the numeral 4 is an even number that can be divided by the numeral 2 with no remainder. Obviously we do not expect the monkeys to understand the concept of even numbers or the operation of division, but it is possible that they understand the order of numerals (their ordinal value) and that the numerals represent specific quantities (their cardinal value). It is also possible that the monkeys do not understand the order of the numerals or the quantities associated with them, but instead respond to the numerals based on a complex matrix of memorized response patterns (e.g., pick the numeral 7 when presented with 6, not when presented with 8) or their hedonic value. The results of the two current studies provided us with a clearer picture of what Arabic numerals mean to the rhesus macaques in our laboratory.

In Experiment 1, the monkeys accurately compared Arabic numerals and analog dot arrays, even on the first exposure to these trials. This indicates that the monkeys were not relying on a complex matrix of learned discriminations. Additionally, the data from Hank suggest that the monkeys were not solving the comparisons based on the hedonic value of the numerals. Although Hank was reinforced proportionally for numerals and not dot displays, his responses to the last half of the novel numeral:dot probes were similar to the responses of the other monkeys, even when numerical value opposed hedonic value. For example, Hank was able to respond correctly to a comparison of four dots and the numeral 3, despite the fact that the numeral 3 had a much richer reinforcement history. The results also indicated that Hank did not base his responses on the probability of reward. He responded correctly on the majority of trials in which an array of five dots was paired with a larger array or numeral, even though he had always been reinforced for choosing arrays of five dots during training. The fact that the monkeys were not responding to numerical pairs based on a complex matrix of memorized responses or hedonic values suggests that these monkeys understood that the Arabic numerals represented absolute values that could be ordered and compared on a relative basis.

Although the results of Experiment 1 suggest that the monkeys were using quantity information to make comparisons between numerals and dot quantities, the quantity information appears to be approximate rather than exact. Performance suffered when the numerical distance between the numeral and dot quantity was small, which is a hallmark of the analog magnitude model of numerical ability (e.g., Brannon & Roitman, 2003; Gibbon, 1977; Meck & Church, 1983). According to this model, numerical performance is based on a continuous representation of magnitude rather a representation of the exact number of items in a set. Memory for the magnitudes associated with each numeral is imperfect so it is more difficult to compare numerals that are close in distance than numerals that are far apart (Dehaene, 1992, 2003; Gallistel & Gelman, 1992, 2000; Whalen, Gallistel, & Gelman, 1999). This model has been used by a number of researchers to explain animal numerical behavior in studies involving analog stimuli (e.g., Beran, 2001, 2004; Beran & Rumbaugh, 2001; Nieder & Miller, 2004)

The results of Experiment 2 provide additional information on the representations underlying the monkeys' use of Arabic numerals. Three of the four animals performed better than chance in a task requiring them to make a number of responses equaling an Arabic numeral. However, on most trials, the monkeys touched a target more times than was represented by the numeral. These data, like the data from Experiment 1, suggest that the monkeys had some understanding of the quantity symbolized by the numerals, but were not representing that quantity precisely. In other words, the monkeys were not enumerating exactly three responses to the numeral 3, but were instead responding in a more approximate manner. Overall, these two studies provide evidence that the rhesus macaques in our laboratory understand the relative values of Arabic numerals and can use this knowledge to compare two numerals. These monkeys also understand that Arabic numerals represent approximate quantity information and can use that information to compare numerals to analog stimuli and to perform a task requiring the enumeration of sequential responses.

Although these numerical abilities are impressive, it is clear that the monkeys do not have a human-like understanding of numerals. Humans use number words and symbols to move beyond the realm of approximation and communicate the precise numerical values required for formal mathematics. The monkeys in this study understood the order of the numerals and could use them to facilitate responding in tasks requiring knowledge of quantity information. In contrast to humans, however, they behaved as if the representations underlying the Arabic numerals were fuzzy approximations of true set size rather than precise quantities. Therefore, what seems to distinguish the symbolic numerical competence of monkeys from that of humans is the representation of exact set sizes across a large range of quantities. Only humans need the exactness of representing numbers such as 9, 13, 142, or even 4 for that matter. Monkeys may need to judge between small sets so that they can make important choices between things like four pieces of food and three pieces, or two predators versus three, but even these judgments do not require exact numerical knowledge, just an ability to distinguish relative numerousness. Outside of laboratory tasks like the ones in this study, monkeys probably never need to know that there are exactly six predators, or to distinguish 16 pieces of fruit from 18, for example. In those cases, the approximate representation of those numbers provides all of the information necessary to aid decision making that increases survival odds. The present findings indicate that, although nonhuman primates do not need to know absolute numerousness in the wild, they can learn symbols that represent such numerousness and use these symbols in a variety of different contexts.

References

Beran, M. J. (2001). Summation and numerousness judgments of sequentially presented sets of items by chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology, 115,* 181-191.

Beran, M. J. (2004). Chimpanzees (*Pan troglodytes*) respond to non-visible sets after one-by-one addition and removal of items. *Journal of Comparative Psychology, 118,* 25-36.

Beran, M. J., Beran, M. M., Harris, E. H., & Washburn, D. A. (2005). Ordinal judgments and summation of nonvisible sets of food items by two chimpanzees (*Pan troglodytes*) and a rhesus macaque (*Macaca mulatta*). *Journal of Experimental Psychology: Animal Behavior Processes, 31,* 351-362.

Beran, M. J., Harris, E. H., Evans, T. A., Klein, E. D., Chan, B., Flemming, T. M., & Washburn, D. A. (2008). Ordinal judgments of symbolic stimuli by capuchin monkeys (*Cebus apella*) and rhesus monkeys (*Macaca mulatta*): The effects of differential and nondifferential reward. *Journal of Comparative Psychology, 122,* 52-61.

Beran, M. J., & Rumbaugh, D. M. (2001). "Constructive" enumeration by chimpanzees (*Pan troglodytes*) on a computerized task. *Animal Cognition, 4,* 81-89.

Biro, D., & Matsuzawa, T. (2001). Use of numerical symbols by the chimpanzee (*Pan troglodytes*): Cardinals, ordinals, and the introduction of zero. *Animal Cognition, 4,* 193-199.

Boysen, S. T., & Berntson, G. G. (1989). Numerical competence in a chimpanzee (*Pan troglodytes*). *Journal of Comparative Psychology, 103,* 23-31.

Boysen, S. T., & Berntson, G. G. (1995). Responses to quantity: Perceptual versus cognitive mechanisms in chimpanzees (*Pan troglodytes*). *Journal of Experimental Psychology:*

Animal Behavior Processes, 21, 82-86.

Brannon, E. M., & Roitman, J. D. (2003). Nonverbal representations of time and number in animals and human infants. In W. H. Meck (Ed.), *Functional and neural mechanisms of interval timing* (pp. 143-182). Boca Raton, FL: CRC Press.

Brannon, E. M., & Terrace, H. S. (1998). Ordering of the numerosities 1–9 by monkeys. *Science, 282*, 746–749.

Brannon, E. M., & Terrace, H. S. (2000). Representation of the numerosities 1-9 by rhesus macaques (*Macaca mulatta*). *Journal of Experimental Psychology: Animal Behavior Processes, 26,* 31–49.

D'Amato, M. R., & Colombo, M. (1988). Representation of serial order in monkeys (*Cebus apella*). *Journal of Experimental Psychology: Animal Behavior Processes, 14,* 131–139.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition, 44,* 1-42.

Dehaene, S. (2003). The neural basis of the Weber-Fechner law: A logarithmic mental number line. *Trends in Cognitive Sciences, 7,* 145-147.

Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition, 44,* 43-74.

Gallistel, C. R., & Gelman, R. (2000). Non-verbal numerical cognition: From reals to integers. *Trends in Cognitive Sciences, 4,* 59-65.

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review, 84,* 279-325.

Harris, E. H., Beran, M. J., & Washburn, D. A. (2007). Ordinal list integration for symbolic, arbitrary, and analog stimuli by rhesus macaques (*Macaca mulatta*). *Journal of General Psychology, 134,* 183-197.

Harris, E. H., & Washburn, D. A. (2005). Macaques' (*Macaca mulatta*) use of numerical cues in maze trials. *Animal Cognition, 8,* 190-199.

Harris, E. H., Washburn, D. A., Beran, M. J., & Sevcik, R. A. (2007). Rhesus monkeys (*Macaca mulatta*) select Arabic numerals or visual quantities corresponding to a number of sequentially completed maze trials. *Learning and Behavior, 35,* 53-59.

Matsuzawa, T. (1985). Use of numbers by a chimpanzee. *Nature, 315,* 57-59.

Meck, W. H., & Church, R. M. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behavior Processes, 9,* 320-324.

Mitchell, R. W., Yao, P., Sherman, P. T., & O'Regan, M. (1985). Discriminative responding of a dolphin (*Tursiops truncatus*) to differentially rewarded stimuli. *Journal of Comparative Psychology, 99,* 218-225.

Murofushi, K. (1997). Numerical matching behavior by a chimpanzee (*Pan troglodytes*): Subitizing and analogue magnitude estimation. *Japanese Psychological Research, 39,* 140-153.

Nieder & Miller (2004). Analog numerical representations in rhesus monkeys: Evidence for parallel processing. *Journal of Cognitive Neuroscience, 16,* 889-901.

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,* 325-339.

Olthof, A., & Roberts, W. A. (2000). Summation of symbols by pigeons (*Columba livia*): The importance of number and mass of reward items. *Journal of Comparative Psychology, 114,* 158-166.

Pepperberg, I. M. (1994). Numerical competence in an African grey parrot (*Psittacus erithacus*).

Journal of Comparative Psychology, 108, 36-44.

Pepperberg, I. M. (2006). Grey parrot (*Psittacus erithacus*) numerical abilities: Addition and further experiments on a zero-like concept. *Journal of Comparative Psychology, 120,* 1-11.

Rumbaugh, D. M., Hopkins, W. D., Washburn, D. A., & Savage-Rumbaugh, E. S. (1989). Lana chimpanzee learns to count by 'numath': A summary of a videotaped experimental report. *Psychological Record, 39,* 459-470.

Rumbaugh, D. M., Richardson, W. K., Washburn, D. A., Savage-Rumbaugh, E. S., & Hopkins, W. D. (1989). Rhesus monkeys (*Macaca mulatta*), video tasks, and implications for stimulus-response spatial contiguity. *Journal of Comparative Psychology, 103,* 32-38.

Terrace, H. S., Son, L. K., & Brannon, E. M. (2003). Serial expertise of rhesus macaques. *Psychological Science, 14,* 66–73.

Washburn, D. A. (1994). Stroop-like effects for monkeys and humans: Processing speed or strength of association? *Psychological Science, 5,* 375–379.

Washburn, D. A., & Rumbaugh, D. M. (1991). Ordinal judgments of numerical symbols by macaques (*Macaca mulatta*). *Psychological Science, 2,* 190–193.

Washburn, D. A., & Rumbaugh, D. M. (1992). Testing primates with joystick-based automated apparatus: Lessons from the Language Research Center's Computerized Test System. *Behavior Research Methods, Instruments, & Computers, 23,* 157-164.

Whalen, J., Gallistel, C. R., & Gelman, R. (1999). Nonverbal counting in humans: The psychophysics of number representation. *Psychological Science, 10,* 130-137.

Xia, L., Emmerton, J., Siemann, M., & Delius, J. D. (2001). Pigeons (*Columba livia*) learn to link numerosities with symbols. *Journal of Comparative Psychology, 115,* 83-91.

Xia, L., Siemann, M., & Delius, J. D. (2000). Matching of numerical symbols with number of responses by pigeons. *Animal Cognition, 3*, 35-43.

Author's Note

This research was supported by a grant (HD-38051) from the National Institute of Child Health and Human Development to the Language Research Center of Georgia State University and research program enhancement support from Georgia State University. The data contained in Experiment 1 of this article were collected as partial fulfillment for the degree of Master of Arts from Georgia State University to Jonathan P. Gulledge. All applicable federal, disciplinary, and institutional rules and regulations regarding animal care and use have been followed in the care and testing of the monkeys. Correspondence concerning this article can be sent to Emily Harris, Language Research Center, Georgia State University, 3401 Panthersville Rd, Decatur, Georgia 30034. Electronic mail may be sent to eharris11@gsu.edu.

Table 2.1

Overall Accuracy Levels across Trial Type for Experiment 1

Trial Type	Mean % Correct	Number of Trials	Std. Deviation
Numeral-Numeral (Training)	84.5*	2399	11.82
Dot-Dot (Training)	89.3*	2385	4.72
Numeral-Dot Probes (Testing)	81.5*	360	2.52

* Performance is significantly better than chance, $p < .01$

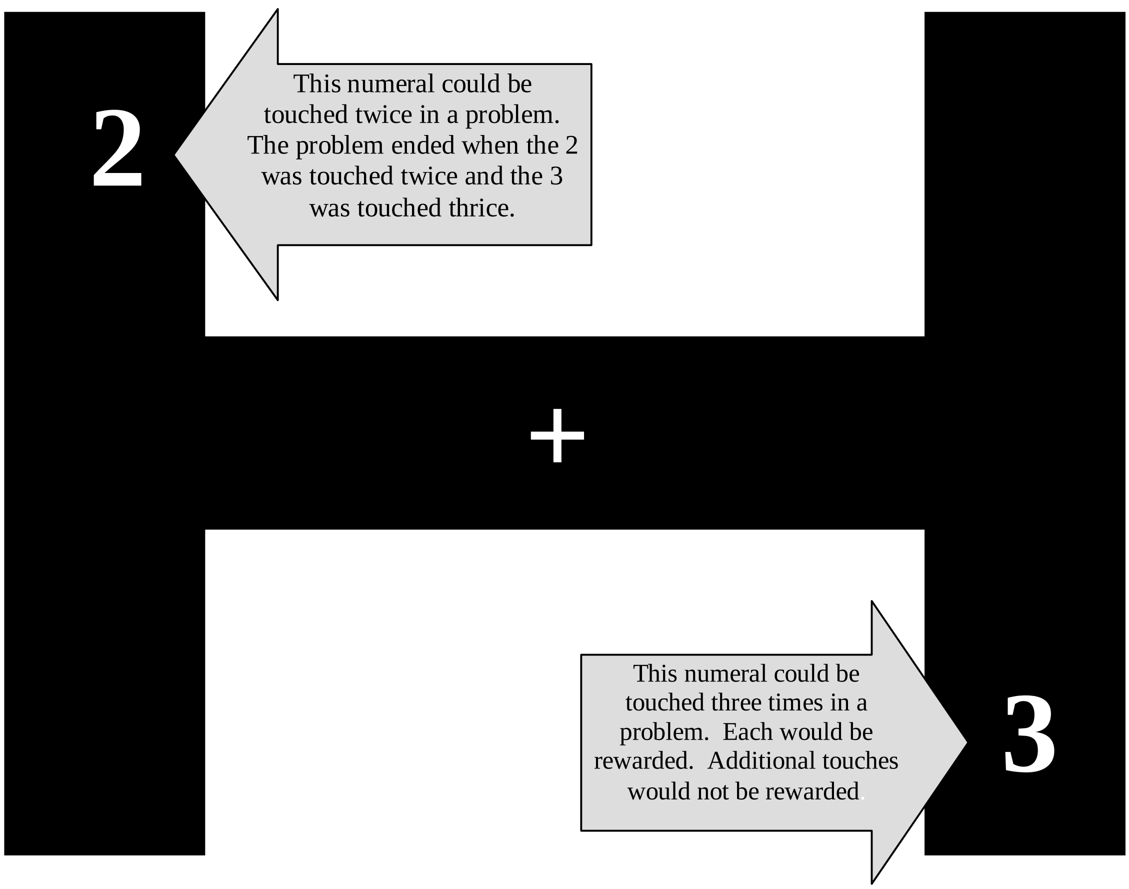
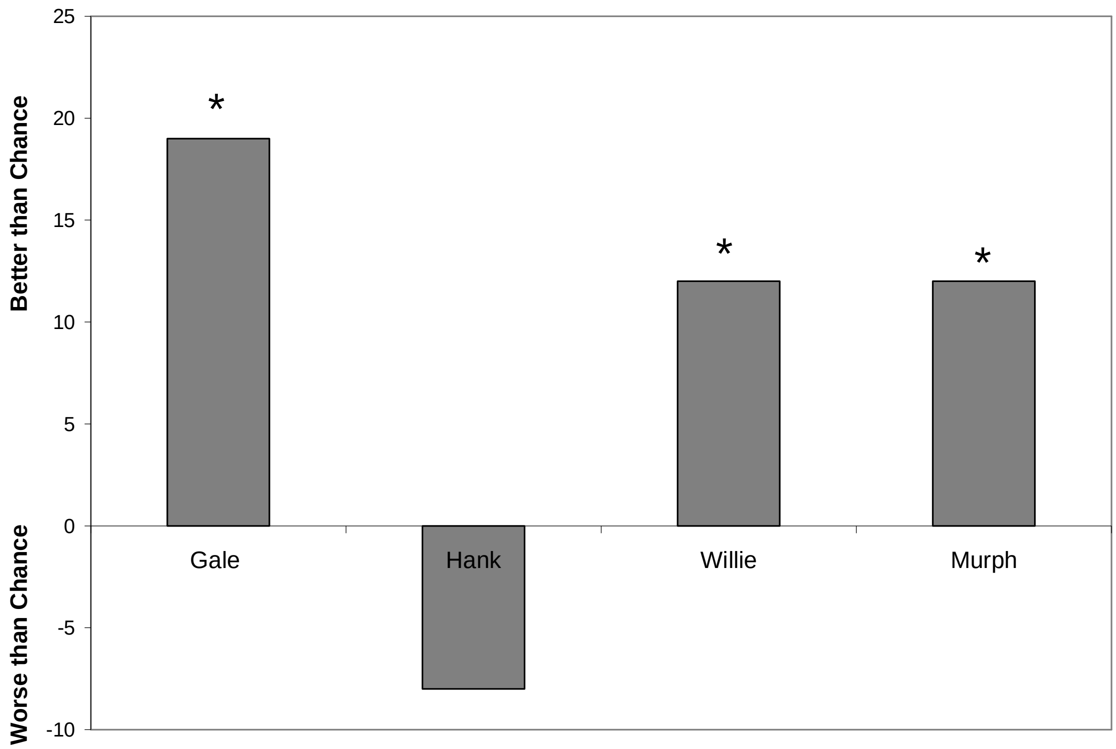


Figure 2.1. Example of the type of display used in Experiment 1 (the arrows and annotations within the arrows did not appear on the monkeys' screens).



* Performance is significantly better than chance, $p < .05$

Figure 2.2. Percent of problems completed without error relative to chance in Experiment 1.

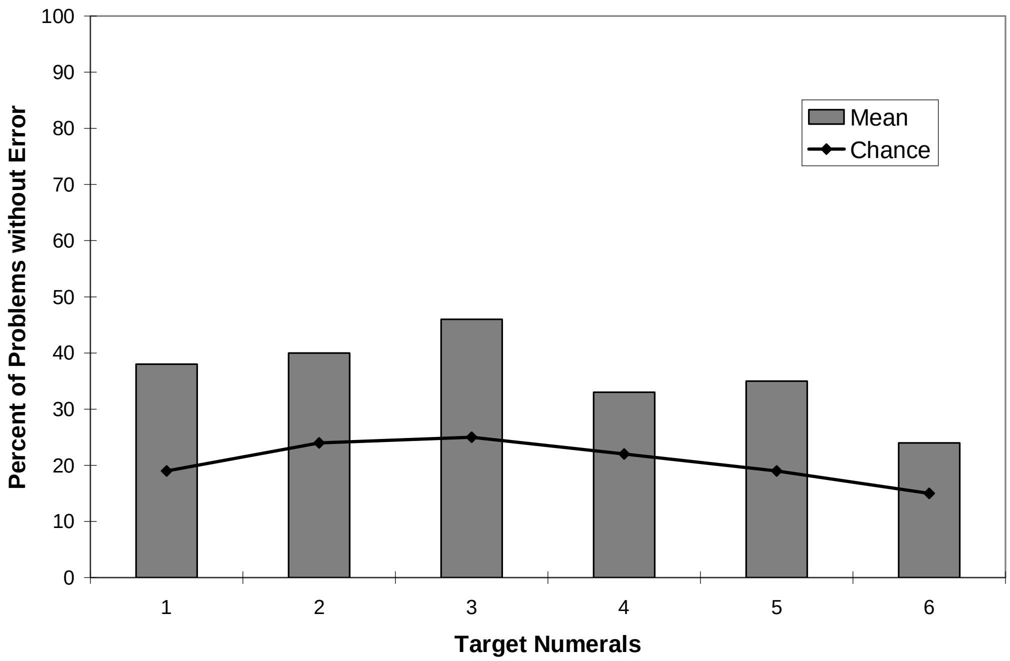


Figure 2.3. Percent of problems completed without error that included a given target numeral for Gale, Willie, and Murph in Experiment 2.

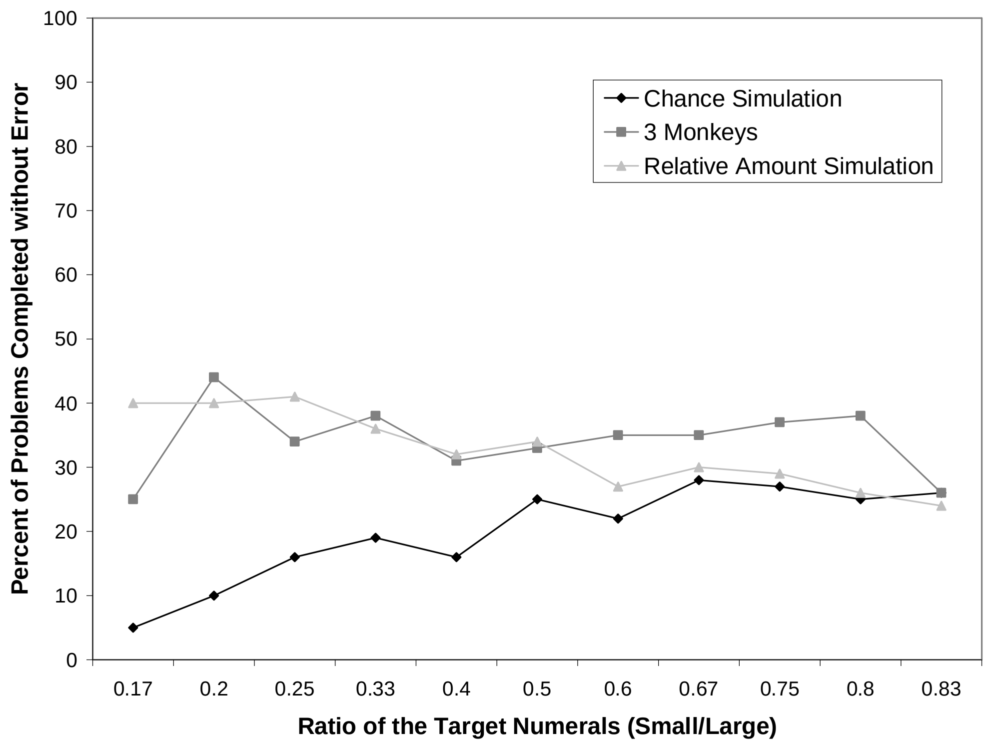


Figure 2.4. Percent of problems completed without error by ratio (small/large target numerals) for Gale, Willie, and Murph in Experiment 2 compared to computer simulations of chance performance and a strategy based on relative amount.

Chapter 3: Ordinal-list integration for symbolic, arbitrary, and analog stimuli by rhesus macaques (*Macaca mulatta*)[2]

Abstract

Two numeral-trained monkeys learned to produce 3 5-item lists of Arabic numerals, colors, and arbitrary signs in the correct sequence. The monkeys then responded at above-chance levels when the authors tested them with nonrewarded pair-wise comparisons of items from different lists, indicating their use of ordinal-position information. The authors also tested the monkeys with nonrewarded pair-wise comparisons of an analog quantity and an item from 1 of the 3 learned lists. Although the monkeys were not trained to serially order analog quantities, 1 monkey correctly integrated the analog quantities with the lists of numerals, colors, and signs. The consistent use of an ordinal rule, despite different types of training and varying degrees of experience with the 4 types of stimuli, suggested that the monkey had a robust concept of ordinality.

---

[2] This chapter was previously published as: Harris, E.H., Beran, M.J., & Washburn, D.A. (2007). Ordinal-list integration for symbolic, arbitrary, and analog stimuli by rhesus macaques (*Macaca mulatta*). *The Journal of General Psychology, 134,* 183-197.

Nonhuman primates can learn to produce serial lists of four, five, and even seven arbitrary stimuli (e.g., D'Amato & Colombo, 1988, 1989; Swartz, Chen, & Terrace, 1991, 2000; Terrace, Son, & Brannon, 2003; Treichler, Raghanti, & Van Tilburg, 2003). In addition to producing such lists, animals also can retain knowledge of several lists at once in long-term memory. This ability allows researchers to investigate the type of knowledge that animals acquire when learning serial lists and the organization that occurs for list items that are presented in different ways (e.g., Treichler & Van Tilburg, 1999, 2002). For example, researchers such as Treichler et al. investigated whether animals infer an integrated serial relationship among items. In the present article, we investigated the question of whether animals encode the ordinal relations between items in the lists.

Early theories of human memory proposed that humans learn a serial list by focusing on associations between adjacent or even remote items in the list. In other words, items are stored as pairs in memory so that each item is associated with another item (Ebbinghaus, 1964; Young, 1961). However, another possibility is that, when mastering a list, humans and animals learn item-position associations. For example, they learn an association between Item 1 and the first ordinal position and between Item 2 and the second ordinal position (e.g., Burns, Dunkman, & Detloff, 1999; Ebenholtz, 1963). Chen, Swartz, and Terrace (1997) reported that rhesus monkeys learned four lists containing four arbitrary items each. The monkeys then learned four 4-item lists that were derived from individual items in the original lists. On two of the new lists, each item's original ordinal position was maintained and, on the other two new lists, the ordinal position of each item was changed. The lists that maintained the ordinal positions were much easier for the monkeys to learn, indicating that the monkeys retained information about the ordinal positions of individual items in each list. Other researchers have used comparisons of two items from different lists to provide evidence that monkeys learn the ordinal positions of each item. For example, monkeys that were trained on a series, $A_1B_1C_1D_1$, and a second series, $A_2B_2C_2D_2$, immediately responded correctly at a greater-than-chance level when researchers presented them with comparisons such as A_1-C_2 and B_1-D_2 (Terrace et al., 2003). Therefore, learning about the ordinal relations of stimuli seemingly emerges for free in the sense that reinforcement contingencies are not tied to ordinal information about list items during training phases. Given that these relations emerged in a number of situations, in the current study we assess the role of previous experience with various types of stimuli on ordinal-list integration.

In the aforementioned studies, the items typically have been arbitrary stimuli, such as photographs, with no meaningful relevance to animals outside of the constraints of the task. For many years, we have presented rhesus monkeys with a variety of stimulus types, some of which have come to operate at or near symbolic levels. For example, monkeys learned to select the larger of two Arabic numerals, and after extensive training with specified pairs, they selected the larger member of a never-before-seen pairing. The monkeys also selected sets of three, four, and even five numerals in descending order (Washburn & Rumbaugh, 1991). The monkeys received a number of food pellets proportional to the numeral or array that they selected (e.g., five pellets for picking the numeral "5" or five dots). Thus, they were always rewarded regardless of their selection. In an extension of Washburn and Rumbaugh's study, Gulledge (1999) demonstrated that monkeys learned to choose the larger of two Arabic numerals or the larger of two arrays of circular dots, and when later presented with a numeral and a dot array, they selected the larger stimulus, even when only correct responses were rewarded. However, in each of those studies, whether the animals responded to numerals on the basis of their ordinal or cardinal value was unknown. If the numerals had ordinal values for the monkeys, each numeral would be linked to all others in a serial list. If the numerals operated with cardinal values, it would be the result of linking each numeral to a specific number of food items.

In the present study, we tested the ability of 2 rhesus monkeys to integrate lists of Arabic numerals, colored squares, arbitrary signs, and analog quantities. Our monkeys had prior testing experience in using Arabic numeral stimuli and analog quantities, for which numerals may have obtained cardinal or ordinal values. Both monkeys also participated in a recent study in which researchers presented them with a series of one to nine computerized maze trials (Harris, Washburn, Beran, & Sevcik, in press). Upon completion of all of the maze trials in a series, the researchers gave the monkeys a same-or-different discrimination involving a numerical stimulus (either an Arabic numeral or a dot quantity) and the letter "D" (for different). Harris et al. rewarded the monkeys for choosing the numerical stimulus if it corresponded to the number of maze trials in the previously completed series. If the numerical stimulus did not match the number of maze trials, they rewarded the monkeys for choosing the "D." In addition to these studies with Arabic numerals, the monkeys also had previous experiences with analog quantities. Both monkeys had participated in a study in which they assessed the number of circular dots on a screen as being either larger or smaller than a learned central value (Beran, Smith, Redford, & Washburn, 2006). Thus, Arabic numerals and analog quantities in the form of circular dots were very familiar stimuli for these monkeys, and this familiarity allowed us to examine spontaneous list integration for meaningful and nonmeaningful stimuli.

Our first experimental question pertained to exactly what numerals meant to the monkeys in terms of how they were represented. If Arabic numerals have ordinal value, the monkeys should learn to produce a list of Arabic numerals faster than they would learn to produce a list of unfamiliar arbitrary stimuli because of both their prior experience with the numerals and their prior knowledge of the ordinal relations between numerals. Conversely, if past experiences using Arabic numerals have led to representations of those numerals that are linked to specific quantities (cardinal value), the monkeys should show no advantage when learning to produce an ordinal list of Arabic numerals because cardinal, and not ordinal, relations were the basis of their earlier judgments (the number of food items received for picking numerals led to a representation of numerals in terms of cardinal value). In addition, these quantity representations should not lead to facilitative effects during integration of the numeral list with lists of arbitrary signs and colors that are only associated with ordinal information. In Experiment 1, monkeys learned to serially order a list of five numerals, a list of five colored squares, and a list of five arbitrary signs. In Experiment 2, the monkeys received nonrewarded pair-wise comparisons of items from different lists, testing the ability of the monkeys to use ordinal position information to integrate the lists.

Our second experimental question pertained to the integration of analog quantities into ordinal lists. Given that these 2 monkeys had previous experience in viewing and responding to a variety of analog stimuli, we investigated whether those types of stimuli could be spontaneously integrated into ordinal lists on the basis of converting the quantity information that was inherent in the analog sets into ordinal information. It was critical that the animals had previous experience with analog sets so that, when the time came to compare a quantity with an item from one of the trained lists, the animals would be familiar with many different analog quantities. In Experiment 3, the monkeys received nonrewarded pair-wise comparisons of analog quantities and items from the three learned lists. We did not train the monkeys to select analog quantities in descending order. We relied on their previous experiences with these types of stimuli to provide them with the necessary information to spontaneously encode the ordinal relations between such stimuli within this new type of task. Although the monkeys had some prior experience in comparing numerals and analog quantities (Gulledge, 1999), they had no experience in comparing colors and analog quantities, comparing signs and analog quantities, or serially ordering lists of analog quantities. Unlike the Gulledge study, in which the monkeys were rewarded every time they selected a numeral or analog quantity, in the current study we provided them with no reward regardless of their selections. In addition, the analog quantities that Gulledge used were uniform circles that did not vary in size. The analog quantities that we used were various polygons ranging in size from 1 to 3 cm that we presented in varying configurations on the screen. Thus, to respond correctly to the pairwise comparisons in Experiment 3, the monkeys would have to apply ordinal information about analog quantities that they obtained from a very different context or use the ordinal information that might spontaneously emerge from extensive experience in viewing and responding to these types of stimuli.

Experiment 1

Method

Subjects. Subjects were 2 male rhesus monkeys (*Macaca mulatta;* called Lou and Murph) aged 11 years. The monkeys were housed individually at the Language Research Center of Georgia State University according to federal animal-housing standards and were not food or water deprived.

Both monkeys had previously been trained to manipulate a joystick and respond to stimuli on a computer screen. For example, in prior tasks investigating numerical ability, both monkeys learned to select the larger of two Arabic numerals (range = 0-9), the larger of two analog dot arrays (range = 1-9), and the larger stimulus when researchers presented them with an

Arabic numeral and dot array (range = 1-9; Gulledge, 1999; Washburn, 1994). Although some monkeys in our laboratory have ordered more than two simultaneously visible Arabic numerals (Washburn & Rumbaugh, 1991), the monkeys involved in the present study had no experience in learning lists or ordering stimuli and were not in visual contact with monkeys performing that type of task.

Apparatus. We tested the monkeys in their home cages using the LRC Computerized Test System (see Rumbaugh, Richardson, Washburn, Savage-Rumbaugh, & Hopkins, 1989, for a description), which consists of a joystick that is attached to a computer and color monitor. The monkeys moved the joystick to control the movement of the cursor on the screen. The computer program, which was written in Visual Basic, recorded the type and number of stimuli that we presented on each trial and the responses that the monkeys made.

Design and Procedure: Phase 1. The procedure that we used was similar to those of previous studies of ordinal knowledge and serial learning in monkeys (e.g., D'Amato & Colombo, 1988; Swartz et al., 1991). At the beginning of each trial, a cursor appeared in the middle of the screen, and the monkey was required to move the cursor and make contact with stimuli positioned randomly around the perimeter of the screen in one of eight possible locations. To prevent the animal from relying on a fixed motor response pattern, the positions of the items varied randomly from trial to trial. After contact with a stimulus, a green border surrounded the stimulus for 300 ms, indicating to the animal that its response had been recorded.

Initially, the monkeys received only the first and second items in a list. We added the next item in the list when a monkey correctly completed 39 of the 60 most recent trials (65% accuracy). The program terminated trials and scored them as incorrect if the monkey skipped an item or made contact with an item that came earlier in the sequence (termed *forward* and *backward errors* by Terrace et al., 2003). We categorized repeated responses to the same stimulus, which occurred on less than 3% of the trials, as backward errors and terminated the trial. Correct trials were rewarded with a melodic sequence of tones and the automatic delivery of a 94-mg fruit-flavored Noyes pellet. Incorrect trials resulted in a negative buzzing tone and a 10 s time-out during which the screen remained black.

We trained the monkeys on three lists: a list consisting of the Arabic numerals "5" through "1" (in descending order), a list consisting of five uniquely colored squares, and a list consisting of five arbitrary symbols ("$," "%," "@," "#," and "*"). During testing, the background was white, and the numerals and signs were black. All stimuli were approximately 5 cm × 5 cm.

After the monkeys reached criterion with the first five-item list, they began training with the first two items on the next list. Lou was trained first on the list of numbers, then on the list of colors, and finally on the list of signs. Murph was trained on the list of colors, then numbers, and finally signs.

Design and Procedure: Phase 2. After reaching criterion with all three of the five item lists, we gave the monkeys a version of the task in which five items were present on each trial, but the type of items (numerals, colors, or signs) varied from trial to trial. The monkeys might be required to order five colors on one trial and five signs on the next trial. The monkeys performed this version of the task in 500-trial sessions until they had achieved 65% accuracy for all three types of stimuli during a session. This phase ensured that the monkeys were still at criterion for all three lists before Experiment 2.

Results and Discussion

Both monkeys learned to produce the three 5-item lists in approximately 25 4-hr sessions. After they had learned all three lists, both monkeys required seven sessions of 500 trials to perform at criterion when the stimulus type varied from trial to trial. The monkeys could attain criterion (39 out of the 60 most recent trials) by responding correctly to 39 consecutive trials. Therefore, 39 trials was the minimum number of trials that we required of the monkeys to satisfy criterion for each training phase with each ordinal list. Table 3.1 shows the actual numbers of trials to criterion that the monkeys required on each phase of training. After the initial two-item phase, the monkeys could have performed the task successfully by executing the known sequence and then responding to the new item. However, they did not appear to be using this list-learning strategy because the number of trials to criterion typically increased as the length of the list increased.

There was no facilitation of list learning when Arabic numerals were the stimuli. One monkey, Murph, showed the greatest difficulty at all set sizes in learning to select Arabic numerals in descending order. Lou had the greatest difficulty with Arabic numerals for two of the set sizes. We had predicted that, if Arabic numerals already had ordinal value for the monkeys (given their previous experiences), learning should occur rapidly. This was especially true because the descending selection order was exactly the same requirement as the optimal responding strategy in their previous number comparison task (Gulledge, 1999). The fact that they did not show better performance, but instead performed poorly with numerals, suggests that those stimuli had not accrued ordinal value during previous exposure.

Because the monkeys used to be able to select the larger of two numerals, one might expect performance to have been higher than what resulted in Experiment 1. However, the most recent experience these monkeys had with numerals was a task in which simply selecting the larger of the two numerals was not an optimal response strategy. Rather, they had to match the number of runs through a maze to the corresponding numeral. Therefore, prior to this training, it did not appear that Arabic numerals carried with them an ordinal value for the monkeys. Although we could not yet state that numerals had cardinal value for the monkeys, we assessed this question in Experiment 2.

Experiment 2

We designed Experiment 2 to determine whether the monkeys learned the ordinal positions of the items in the trained lists of colors, numerals, and signs. Although we continued to present the monkeys with trials like those at the end of Experiment 1, we also introduced nonrewarded probe trials in which we presented two items from different lists and different ordinal positions within the learned lists. If the monkeys had learned the ordinal locations of the various stimuli in the learned lists, they should have selected the correct stimuli at levels above those of chance. Such results would replicate earlier studies that showed ordinal learning by nonhuman animals (e.g., Swartz et al., 2000; Terrace et al., 2003; Treichler et al., 2003) and provide the foundation for Experiment 3, in which we would examine spontaneous ordinal-list integration for unlearned analog stimuli.

Method

Subjects and Apparatus. In Experiment 2, we used the same subjects and apparatus from Experiment 1.

Design and Procedure. Experiment 2 consisted of sequencing trials that were identical to those in Phase 2 of Experiment 1 and pair-wise comparison trials. The pair-wise comparison trials, which occurred on every 5th trial, consisted of randomly chosen items from two different lists that occupied different ordinal positions. For example, a comparison trial might consist of a black square (the first item in the color list) and the "@" sign (the third item in the sign list). The two items were side by side on the screen, and the computer program randomly assigned one item to the left side and one to the right. After subjects selected one of the items, the comparison trial ended, and the computer presented a new five-item sequence trial. We provided no positive or negative feedback for the comparison trials, although feedback in the form of tones, pellets, and time-outs continued to be provided for the sequencing trials. The monkeys performed this task twice a week for 4-hr sessions until they had completed 500 comparison trials.

Results and Discussion

We grouped together comparisons involving a color and a numeral, comparisons involving a numeral and a sign, and comparisons involving a color and a sign. We considered a trial correct if the monkey selected the stimulus with the lower ordinal position. For example, if the trial involved a black square (the first item in the color list) and the numeral "4" (the second item in the numeral list), the correct response would be the black square. For all comparison types (colors and numerals, colors and signs, numerals and signs), performance was significantly above chance levels ($p < .05$), according to a sign test that compared performance with a 50% chance level of performance (Figure 3.1).

As in other serial learning studies (e.g., Terrace et al., 2003), we found a distance effect for the comparison trials in which comparisons between items from more disparate ordinal positions on different lists were easier for the monkeys. Accuracy increased as the ordinal distance between the two probe stimuli increased. Even with a small range of distances, correlations between distance and accuracy were very high for both monkeys - Murph: $r(2) = .84$, $p = .08$; Lou: $r(2) = .93$, $p < .05$.

The accuracy for the monkeys on the pair-wise comparison trials indicates that they were responding on the basis of the ordinal position of the items on these three learned lists. The fact that accuracy increased as the ordinal distance between the two stimuli increased indicates that this ordinal knowledge is probably inexact (i.e., the monkeys know that the numeral "4" is near the beginning of the number list, but they may not know that it occupies the second ordinal position).

Experiment 3

Typically, list items in these types of experiments are arbitrary, single stimuli that inhabit ordinal locations in a sequence. Of course, the subject must learn the ordinal value of a stimulus because nothing inherent in the stimulus itself or in its relation to another stimulus denotes its ordinal position in the list. However, other stimuli do provide inherent ordinal information when compared with each other, if nonhuman animals have an understanding and a responsiveness to numerical properties of a stimulus set. When analog quantities are shown, nonhuman animals respond to their ordinal relations on the basis of ascending or descending numerosity, and not specific stimulus properties such as color, arrangement, or size (e.g., Brannon & Terrace, 2000; Emmerton, Lohmann, & Niemann, 1997; Judge, Evans, & Vyas, 2005; Smith, Piel, & Candland, 2003; Thomas & Chase, 1980). The question that we addressed in Experiment 3 was whether monkeys encode such analog quantities in terms of their ordinal position as determined by their numerosity when they are presented in comparison with unitary stimuli from learned ordinal lists. If so, monkeys should select the analog stimulus set when its numerosity exceeds the learned ordinal position of the unitary comparison stimulus, whereas they should select the unitary stimulus set when its learned ordinal position exceeds that of the analog quantity.

It is important to note that the monkeys' previous experience with analog stimuli (e.g., Gulledge, 1999) was critical to our ability to ask this question. Because these 2 monkeys had compared two sets of dots and compared a set of dots with a numeral, they had ample opportunity to learn about the relation between differing numbers of analog stimuli. They ultimately responded appropriately in selecting the correct stimulus with those comparisons. However, we do not know whether the monkeys were encoding the number of items in the analog set and comparing that with the representation of the cardinal value of the Arabic numeral member of the pair. They might simply have learned which analog set sizes were correct choices in combination with some numbers but incorrect choices in combination with others. If that were true, the monkeys would not be able correctly to select the larger of an analog set of stimuli and either a color or sign from those newly learned lists. However, if the monkeys previously learned about the ordinal positions of analog stimuli from their cardinal values (i.e., their actual quantitative properties), they should be able to choose correctly no matter which list we paired with an analog quantity.

Method

Subjects and Apparatus. In Experiment 3, we used the same subjects and apparatus from Experiments 1 and 2.

Design and Procedure. The analog quantities were groups of up to five black polygons (squares, rectangles, circles, ovals, parallelograms) with heights and widths ranging from approximately 1 cm to 3 cm. The computer randomly selected the polygons and placed each within one of 81 locations in a 9×9 matrix. The positions of the polygons changed from trial to trial. These stimuli were completely novel and did not resemble the analog stimuli that Gulledge (1999) previously presented to the monkeys in a pair-wise comparison task.

Because we were interested in whether analog quantities spontaneously would accrue ordinal value within the constraints of our task, we offered the monkeys no opportunity to respond to those stimuli prior to the nonrewarded probe trials. In contrast to numerals, colors, and signs, the monkeys never learned to select analog sets in descending order and were never exposed to these analog sets before the first probe trial. Baseline trials only involved five colors, numerals, or signs being presented on the screen, never analog quantities.

The computer program presented pair-wise comparison trials on every 5th trial as during Experiment 2. In between pair-wise comparison trials, the monkeys received the same sequencing trials as in Experiment 2 (with colors, numerals, or signs, but never analog quantities). There were six possible pair-wise comparisons: a numeral and a color, a numeral and a sign, a color and a sign, an analog quantity and a numeral, an analog quantity and a color, and an analog quantity and a sign. The computer program randomly chose pair-wise comparison stimuli from two different lists, with the constraint that they had to be in two different ordinal positions, and placed them on opposite sides of the screen. We provided no positive or negative feedback for any responses during the probe trials. The monkeys performed this task twice a week for 4-hr sessions until they had completed 500 pair-wise comparison trials.

Results and Discussion

We considered a trial correct if the monkey selected the stimulus with the lower ordinal position. For example, if the trial involved a black square (the first item in the color list) and four items in an analog set (the second ordinal position in the analog list), the correct response would be the black square. We analyzed the pair-wise comparison trials the same way as in Experiment 2. For all six comparisons (including those with analog quantities), Murph's performance was significantly above chance ($p < .05$) according to a sign test. However, Lou exceeded chance ($p < .05$) only for comparisons between colors and analog quantities (Figure 3.2).

As in Experiment 2, accuracy increased as the ordinal distance between the two comparison stimuli increased (Murph: $r(2) = .91$, $p < .05$; Lou: $r(2) = .93$, $p < .05$). An important aspect of this distance effect is that it allowed us to take a closer look at the performance of Murph on probe trials with analog quantities. Although Murph's performance was statistically better than chance in comparing analog quantities with items from the other trained lists, one could argue that such performance could emerge not on the basis of his incorporating ordinal information inherent in the analog quantities, but on the basis of previous experience in selecting larger sets of analog quantities over smaller sets of analog quantities. In such a case, Murph should have shown a bias in selecting larger analog sets as compared with smaller analog sets, independent of the comparison stimulus. Given that the monkeys' selection of larger analog sets, as compared with smaller analog sets, also indicated competence in using ordinal information in these probe trials, we could not simply report the frequencies of selecting different analog quantities. However, we could look at shifts in the likelihood of selecting the most extreme analog quantities as a function of the difference between those analog quantities and the comparison stimuli. When the computer program presented a single polygon with list items from Ordinal Position 4 (the second-to-last position), subjects selected the single polygon on 87% of the trials. Subjects selected single analog quantities approximately 50% of the time when they were presented with list items from Ordinal Position 3. If Murph learned through previous experience only that single polygons were not good stimuli to select, he should have shown little or no responding to those stimuli, and he certainly would not have shown preferential responding to single polygons when the comparison stimulus was from a close ordinal position.

In addition, although performance was very high whenever the computer program presented five polygons, the occasional incorrect selection of a trained list stimulus instead of the five polygons always occurred when that comparison stimulus was from Ordinal Position 2 in its trained list (providing a distance of only one ordinal position). When we looked only at comparisons of analog quantities to learned list stimuli, Murph still exhibited a very high correlation between performance and the ordinal distance between stimuli being compared, $r(2) = .97$, $p < .05$.

General Discussion

As we predicted, both monkeys learned to produce three 5-item lists at greater-than-chance levels. During the initial presentation of each list, only two items were present, and we added a new item each time the monkeys reached an accuracy criterion. This successive method has been effective in training other rhesus monkeys to produce ordinal lists (e.g., D'Amato & Colombo, 1998; Swartz et al., 1991), and we have successfully replicated those reports.

Previously, we have trained rhesus monkeys to select the larger of two Arabic numerals or the larger member of a pair containing a numeral and a dot quantity. Although the monkeys could have learned something about the approximate (or exact) number of items associated with those numerals, they might also have simply learned a number of ordinal pairings independent of numerosity. If the latter were true, and if the subjects learned the ordinal values of the numerals, we would have expected the monkeys to reach training criterion more quickly with numerals than with either of the two novel stimulus sets. However, this was not the case, suggesting that either (a) ordinal knowledge did not inhere in the representations formed during the presentation of Arabic numerals in previous studies or (b) if it did, such learning was lost as a result of subsequent studies in which numerals took on cardinal values (e.g., Harris et al., in press). After ordinal training in the present study, the monkeys performed significantly above chance for all types of pair-wise comparisons with items from different lists. Such results indicate that the monkeys were responding on the basis of the newly learned ordinal positions of the items.

Perhaps the more exciting finding is that, in addition to successfully comparing numerals, colors, and signs that were trained as serial lists, one monkey was able to make ordinal comparisons using analog quantities. Murph performed above chance levels for all comparison types, including those that involved analog quantities. As we previously mentioned, both monkeys had prior experience comparing analog dot quantities with Arabic numerals (Gulledge, 1999), but they had received no serial training involving lists of quantities. However, when we presented analog quantities within the context of making ordinal judgments, one monkey spontaneously used the magnitude of the polygon set to determine its ordinal position relative to the learned-list stimuli. Although the monkeys had prior experience performing pair-wise comparisons with analog dot quantities and Arabic numerals (Gulledge), the analog quantities in that experiment were uniform dots and the monkeys were rewarded regardless of selection. In contrast, the analog quantities in the present experiment were polygons that varied in size, and we gave no reward on pair-wise comparison trials. Therefore, applying ordinal information from Gulledge's previous study to the current task indicates impressive generalization abilities. In addition, Murph integrated the analog sets into lists of colors and signs, stimuli that were never previously paired together. Thus, for Murph, analog stimuli took on both cardinal and ordinal values within specified tasks.

Although researchers have demonstrated that monkeys can respond to analog quantities on the basis of ascending or descending numerosity (Brannon & Terrace, 2000; Emmerton et al., 1997; Judge et al., 2005; Smith et al., 2003; Thomas & Chase, 1980), the present study is the first evidence that a monkey will respond on the basis of ordinal position on nonrewarded trials in which an analog quantity is presented in comparison with a unitary stimulus from a learned ordinal list. Unfortunately, only one of the two monkeys performed at a high level in Experiment 3. Although Lou exceeded the chance levels for the first set of probe trials (Experiment 2), his performance was lower than that of Murph, and he did not sustain above-chance performance during the second set of probe trials, including those with the analog stimuli (Experiment 3). This individual difference is important in illustrating the fragile nature of these representations. However, it is notable that the one comparison in which Lou performed above chance involved colored squares that he was trained to order sequentially and analog quantities that he was never trained to order sequentially, suggesting that he was able to apply an ordinal rule with analog stimuli, despite the different types of training with colors and analog quantities.

Conclusion

Both monkeys showed evidence of ordinal knowledge when we tested them with arbitrary stimuli and familiar Arabic numerals that they had learned to sequence and that had previously been associated with quantity information. One monkey also incorporated ordinal information that was inherent in analog dot quantities. The monkeys' ability to respond to the ordinal information of these different stimuli suggests that they spontaneously attended to ordinal position. These data, combined with the data of other studies (e.g., Brannon & Terrace, 2000; Chen et al., 1997; Terrace et al., 2003), provide a glimpse of what is perhaps a broad concept of ordinal position in rhesus monkeys that is readily gleaned from various properties of stimulus sets. If this concept is true, future researchers could design studies to examine monkeys' immediate incorporation of such ordinal information in cross-list comparisons of different stimulus properties, such as lists of larger and smaller stimuli compared with lists of brighter and dimmer stimuli.

References

Beran, M. J., Smith, J. D., Redford, J. S., & Washburn, D. A. (2006). Rhesus macaques (*Macaca mulatta*) monitor uncertainty during numerosity judgments. *Journal of Experimental Psychology: Animal Behavior Processes, 32,* 111-119.

Brannon, E. M., & Terrace, H. S. (2000). Representation of the numerosities 1-9 by rhesus macaques *(Macaca mulatta). Journal of Experimental Psychology: Animal Behavior Processes, 26,* 31-49.

Burns, R. A., Dunkman, J. A., Jr., & Detloff, S. L. (1999). Ordinal position in the serial learning of rats. *Animal Learning and Behavior, 27,* 272-279.

Chen, S., Swartz, K. B., & Terrace, H. S. (1997). Knowledge of the ordinal position of list items in rhesus monkeys. *Psychological Science, 8,* 80-86.

D'Amato, M. R., & Colombo, M. (1988). Representation of serial order in monkeys *(Cebus apella). Journal of Experimental Psychology: Animal Behavior Processes, 14,* 131-139.

D'Amato, M. R., & Colombo, M. (1989). Serial learning with wild card items by monkeys (*Cebus apella*): Implications for knowledge of ordinal position. *Journal of Comparative Psychology, 103,* 252-261.

Ebbinghaus, H. (1964). Memory: A contribution to experimental psychology. Oxford, England: Dover. (Original work published 1885)

Ebenholtz, S. M. (1963). Serial learning: Position learning and sequential associations. *Journal of Experimental Psychology, 66,* 353-362.

Emmerton, J., Lohmann, A., & Niemann, J. (1997). Pigeons' serial ordering of numerosity with visual arrays. *Animal Learning and Behavior, 25,* 234-244.

Gulledge, J. P. (1999). Judgments of dot arrays and Arabic numerals by rhesus macaques:

Evidence of concept of number. Unpublished master's thesis, Georgia State University, Atlanta.

Harris, E. H., Washburn, D. A., Beran, M. J., & Sevcik, R. A. (in press). Rhesus monkeys (*Macaca mulatta*) select Arabic numerals or visual quantities corresponding to a number of sequentially completed maze trials. *Learning and Behavior*.

Judge, P. G., Evans, T. A., & Vyas, D. K. (2005). Ordinal representation of numeric quantities by brown capuchin monkeys (*Cebus apella*). *Journal of Experimental Psychology: Animal Behavior Processes, 31,* 79-94.

Rumbaugh, D. M., Richardson, W. K., Washburn, D. A., Savage-Rumbaugh, E. S., & Hopkins, W. D. (1989). Rhesus monkeys (*Macaca mulatta*), video tasks, and implications for stimulus-response spatial contiguity. *Journal of Comparative Psychology, 103,* 32-38.

Smith, B. R., Piel, A. K., & Candland, D. K. (2003). Numerity of a socially housed hamadryas baboon (*Papio hamadryas*) and a socially housed squirrel monkey (*Saimiri sciureus).* *Journal of Comparative Psychology, 117,* 217-225.

Swartz, K. B., Chen, S., & Terrace, H. S. (1991). Serial learning by rhesus monkeys: I. Acquisition and retention of multiple four-item lists. *Journal of Experimental Psychology: Animal Behavior Processes, 17,* 396-410.

Swartz, K. B., Chen, S., & Terrace, H. S. (2000). Serial learning by rhesus monkeys: II. Learning four-item lists by trial and error. *Journal of Experimental Psychology: Animal Behavior Processes, 26*, 274-285.

Terrace, H. S., Son, L. K., & Brannon, E. M. (2003). Serial expertise of rhesus macaques. *Psychological Science, 14,* 66-73.

Thomas, R. K., & Chase, L. (1980). Relative numerousness judgments by squirrel monkeys.

Bulletin of the Psychonomic Society, 16, 79-82.

Treichler, F. R., Raghanti, M. A., & Van Tilburg, D. N. (2003). Linking of serially ordered lists by macaque monkeys (*Macaca mulatta*): List position influences. *Journal of Experimental Psychology: Animal Behavior Processes, 29,* 211-221.

Treichler, F. R., & Van Tilburg, D. (1999). Training requirements and retention characteristics of serial list organization by macaque monkeys. *Animal Cognition, 2,* 235-244.

Treichler, F. R., & Van Tilburg, D. (2002). Premise-pair training for valid tests of serial list organization in macaques. *Animal Cognition, 5,* 97-105.

Washburn, D. A. (1994). Stroop-like effects for monkeys and humans: Processing speed or strength of association? *Psychological Science, 5,* 375-379.

Washburn, D. A., & Rumbaugh, D. M. (1991). Ordinal judgments of numerical symbols by macaques *(Macaca mulatta). Psychological Science, 2,* 190-193.

Young, R. K. (1961). The stimulus in serial verbal learning. *American Journal of Psychology, 74,* 517-528.

Author Note

This research was supported by Grant HD–38051 from the National Institute of Child Health and Human Development to the Language Research Center of Georgia State University, a Research Enhancement grant from Georgia State University, and a Rumbaugh Fellowship to E. H. Harris. All applicable federal, disciplinary, and institutional rules and regulations regarding animal care and use have been followed in the care and testing of the monkeys. Address correspondence to Emily H. Harris, Language Research Center, Georgia State University, 3401 Panthersville Road, Decatur, GA 30034; eharris11@gsu.edu (e-mail).

Table 3.1

The Number of Trials to Criterion for Each Phase of Training

Lou	2 Stimuli	3 Stimuli	4 Stimuli	5 Stimuli
List 1 - Numbers	85	516	1,306	1,649
List 2 - Colors	47	2,025	2,166	123
List 3 - Signs	66	60	171	464

Murph	2 Stimuli	3 Stimuli	4 Stimuli	5 Stimuli
List 1 - Colors	48	55	172	1,918
List 2 - Numbers	394	1,626	2,495	3,998
List 3 - Signs	54	72	420	1,269

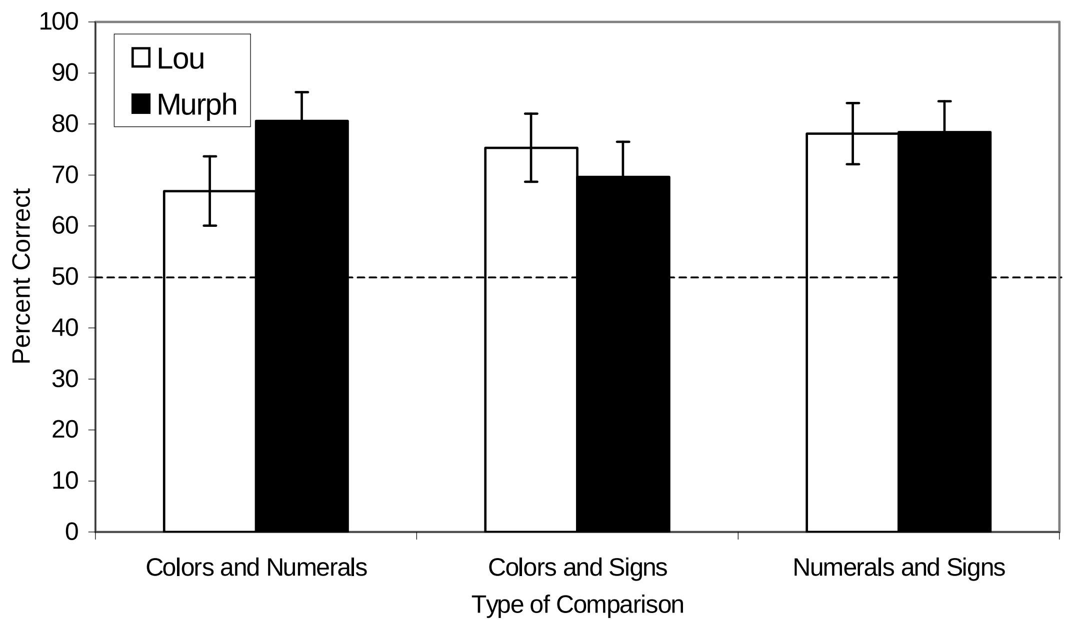


Figure 3.1. Performance on the three types of probe trials from the trained lists. Bars indicate 95% confidence intervals for proportions. Dotted line denotes chance.

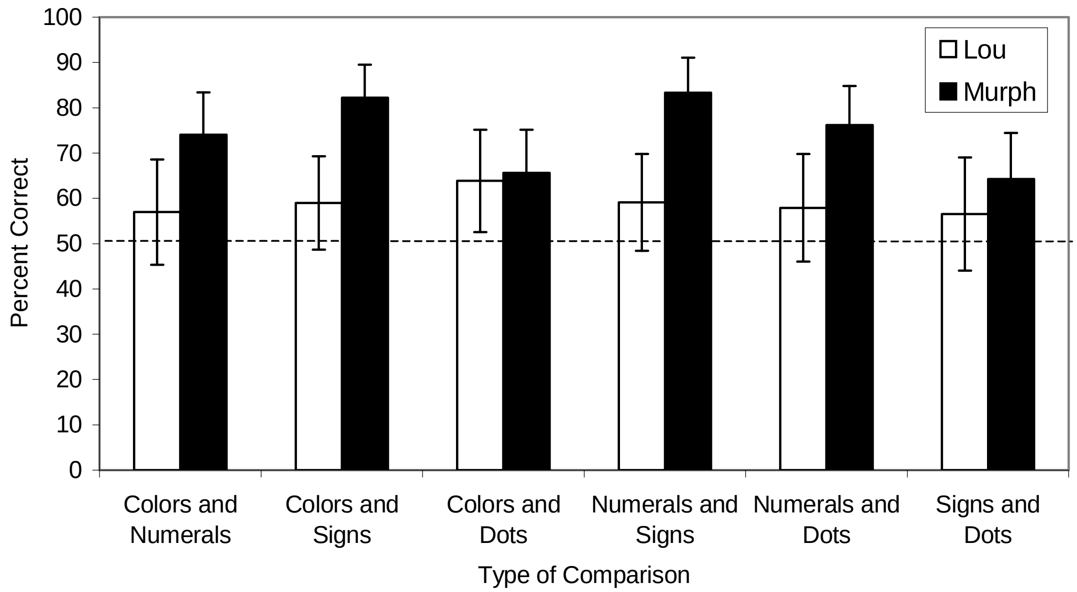


Figure 3.2. Performance on the six types of probe trials from the trained lists and the analog dot quantities. Bars indicate 95% confidence intervals for proportions. Dotted line denotes chance.

Chapter 4: Macaques' (*Macaca mulatta*) use of numerical cues in maze trials[3]

Abstract

We tested the ability of number-trained rhesus monkeys to use Arabic numeral cues to discriminate between different series of maze trials and anticipate the final trial in each series. The monkeys' prior experience with numerals also allowed us to investigate spontaneous transfer between series. A total of four monkeys were tested in two experiments. In both experiments, the monkeys were trained on a computerized task consisting of three reinforced maze trials followed by one nonreinforced trial. The goal of the maze was an Arabic numeral 3, which corresponded to the number of reinforced maze trials in the series. In experiment 1 (n=2), the monkeys were given probe trials of the numerals 2 and 4 and in experiment 2 (n=2), they were given probe trials of the numerals 2–8. The monkeys receiving the probe trials 2 and 4 showed some generalization to the new numerals and developed a pattern of performing more slowly on the nonreinforced trial than the reinforced trial before it for most series, indicating the use of the changing numeral cues to anticipate the nonreinforced trial. The monkeys receiving probe trials of the numerals 2–8 did not predict precisely when the nonreinforced trial would occur in each series, but they did incorporate the changing numerals into their strategy for performing the task. This study provides the first evidence that number-trained monkeys can use Arabic numerals to perform a task involving sequential presentations.

---

[3] This chapter was previously published as: Harris, E.H., & Washburn, D.A. (2005). Macaques' (*Macaca mulatta*) use of numerical cues in maze trials. *Animal Cognition, 8,* 190-199.

Many species of animals including pigeons, rats, raccoons, salamanders, monkeys, and chimpanzees have exhibited some form of numerical knowledge either spontaneously or with training (e.g., Beran and Rumbaugh 2001; Boysen and Berntson 1989; Brannon and Terrace 1998; Capaldi and Miller 1988; Emmerton 1998; Hauser et al. 2000; Matsuzawa 1985; Uller et al. 2003). Most numerical studies with animals have focused on counting behavior (e.g., Davis and Bradford 1986; Beran and Rumbaugh 2001) or relative numerousness judgments. Relative numerousness judgments have typically involved the comparison between visible quantities (e.g., Brannon and Terrace 1998, 2000; Rumbaugh et al. 1987; Thomas et al. 1980; Uller et al. 2003) or visible symbols that represent quantities (e.g., Olthof et al. 1997; Washburn and Rumbaugh 1991).

In a few studies, however, animals have been required to respond to the numerousness of stimuli or sequences that were not simultaneously visible. Rhesus monkeys (Hauser et al. 2000), chimpanzees (Beran 2001), and orangutans (Call 2000) have shown the ability to watch food items placed sequentially into opaque containers, and subsequently to select the container with the most items. Cotton-top tamarins have been shown to discriminate between the number of syllables in two sequences of speech, even while continuous variables such as sequence duration, item duration, inter-stimulus interval, and overall energy were controlled (Hauser et al. 2003).

However, the capacity to respond to nonvisible numerousness is not limited to primates. Capaldi and Miller (1988) used the sequential presentation of events to investigate the ability of rats to count reinforced maze trials. They presented rats with either three or four maze trials. The three-trial series consisted of two reinforced trials followed by a nonreinforced trial (RRN) and the four-trial series consisted of one nonreinforced trial followed by two reinforced trials and a nonreinforced trial (NRRN). Results indicated that in both series the rats ran more slowly on the terminal N trial than any other trial, even when confounding temporal and odor cues were controlled. Capaldi and Miller concluded that the rats were counting the reinforced trials and using that numerical cue to predict when the nonreinforced trial would occur.

These findings have been replicated and extended numerous times (e.g., Burns and Criddle 2001; Burns et al. 2004; Capaldi and Miller 2004). In one of these subsequent studies, Burns et al. (1995) systematically varied the inter-trial intervals in a series of runway trials from 20 to 120 s and obtained results similar to those of Capaldi and Miller (1988). Due to the large variation in inter-trial intervals, Burns and colleagues concluded that the slower running times observed on the terminal N trials could not be explained by rhythmic cues, as had been suggested by Davis and Pérusse (1988). The use of rhythmic cues is an extension of simultaneous subitizing (Mandler and Shebo 1982; Piazza et al. 2002) that applies to sequentially presented items or events.

Rats have also been trained to discriminate between different series of reinforced and nonreinforced trials using brightness and texture cues on the runway floor. For instance, Burns et al. (1999) consistently presented rats with a rough and white floor during an XNY series (where the X and Y represented different food items) and a smooth and black floor during a ZNN series (where the Z represented a third type of food item). Using this procedure, the researchers were able to compare performance between more than one series in the same group of rats. For both series, the rats developed faster running for rewarded trials than for nonrewarded trials.

Monkeys previously trained to make ordinal judgments using Arabic numerals provide a unique opportunity to study the use of numerical cues and spontaneous transfer between series. Arabic numerals, instead of the texture of runway floors, can be used as a cue to help the monkeys determine which type of series is being presented. This in turn, could act as a cue to help them predict when a nonreinforced trial will occur. Because the monkeys should not require additional training on what the Arabic numerals mean, they might then show flexibility in anticipating the identity of nonreinforced trials that occur at various places in a maze sequence.

The monkeys involved in the current study previously learned to select the larger of two Arabic numerals (0-9) to receive the corresponding number of food pellets (Washburn and Rumbaugh 1991). They also learned to select the larger of two analog quantities, such as arrays of letters (Washburn 1994), and have demonstrated the ability to choose the larger stimulus at a greater than chance level when presented with one analog quantity (such as dots) and one Arabic numeral (Gulledge 1999). Based on the monkeys' success in previous numerical tasks involving the simultaneous presentation of analog quantities and numerals we hypothesized that the monkeys would use Arabic numerals to perform a task involving the sequential presentation of maze trials. More specifically, we hypothesized that the monkeys would use the numerals to discriminate among different series of reinforced and nonreinforced computerized maze trials and anticipate the nonreinforced trials. We also hypothesized that the monkeys' prior knowledge of Arabic numerals would allow for spontaneous transfer from one Arabic numeral to another during this sequential task. To test these hypotheses, we trained all of the monkeys on an RRRN series and then introduced probe series involving different numbers of reinforced trials.

Experiment 1

In experiment 1, two rhesus monkeys were trained on a computerized maze series consisting of three reinforced trials followed by one nonreinforced trial (an RRRN series). The numeral 3 was used as the goal of the maze and acted as a cue to the number of reinforced trials that would occur before the nonreinforced trial. After the monkeys had developed a pattern of performing more slowly on the nonreinforced trial in each series compared to the reinforced trial before it, they were introduced to probe trials consisting of the numerals 2 and 4 (an RRN and RRRRN series). The goal was to assess the ability of the monkeys to use the changing target numeral to predict when the nonreinforced trial would occur.

Method

Subjects. Two male rhesus monkeys (*Macaca mulatta*) participated in this study. The monkeys, Murph and Lou, were both 10 years old and had participated in several previous studies that required them to make ordinal judgments using Arabic numerals (following the methods described by Washburn and Rumbaugh 1991). They also had participated in numerous computerized joystick tasks related to various other areas of cognitive research (e.g., Smith et al. 2003; Washburn and Gulledge 2002; Washburn and Rumbaugh 1997). The monkeys were individually housed according to federal animal housing standards and were not food or water deprived during this study.

Apparatus. The monkeys were tested in their home cages using the LRC Computerized Test System (see Rumbaugh et al.1989, for a description) consisting of a joystick attached to a Compaq computer and 17-inch color monitor. The monkeys moved the joystick to control the movement of a cursor on the screen. The computer program recorded the stimuli that were presented along with the amount of elapsed time before the monkey initiated the start of the trial and the amount of time required to complete the trial. Pellets were dispensed automatically upon completion of reinforced trials.

Task. The computerized display consisted of a black H-shaped maze, approximately 21 cm × 29 cm, on a white background (see Figure 4.1). A computer-generated white Arabic numeral, approximately 2.5 cm × 1.8 cm appeared in the upper left hand corner of the maze. The monkeys moved the joystick to begin the trial and to begin measurement of response time. At the beginning of each trial the cursor appeared in the lower right hand corner of the maze and the monkeys were required to move the cursor through the maze and make contact with the Arabic numeral in order to complete the trial successfully.

Training procedure. The monkeys were trained using only the Arabic numeral 3 in the display. The numeral 3 corresponded to a series consisting of three reinforced trials followed by one nonreinforced trial (RRRN). The inter-trial interval was 5 s, during which the screen remained black. Upon successful completion of the reinforced trials there was sound feedback and the automatic delivery of a 97-mg fruit-flavored Noyes pellet. No feedback was given for completion of the nonreinforced trials and the screen remained black for 15 s before a new series began.

After several sessions of training the data were analyzed to determine whether the monkeys should be moved to test phase or should continue training using only the Arabic numeral 3. Monkeys were considered ready for the test phase if they showed significantly slower response times on the fourth (nonreinforced) trials compared to the third (reinforced) trials. Lou received 1,000 training trials (250 series) and Murph received 1,300 training trials (325 series) over the course of two sessions before they were moved to the test phase of the experiment.

Testing procedure. During testing, probe trials consisting of the Arabic numerals 2 and 4 were randomly interspersed with the familiar Arabic numeral 3. As was true during training, the Arabic numeral displayed during testing corresponded to the number of reinforced trials that would occur before one nonreinforced trial. For example, the target numeral 4 corresponded to an RRRRN series. Both monkeys received a total of 1,500 test trials over the course of two sessions.

Results

Both start times and response times were recorded for each trial and those measures were used to compute run time. Start time was defined as the amount of elapsed time from the time at which the maze appeared on the screen to the time at which the monkey initiated the trial by moving the joystick. Response time was defined as the time required to complete the trial successfully once it had been initiated. Run time is the sum of start time and response time for each trial. Analyses were completed separately for start time, response time, and run time. Start time was found to be fairly constant, causing response time and run time to be highly correlated, $r(5{,}300) = 0.85$, $p < 0.01$. Start time and run time were not as highly correlated, $r(5{,}300) = 0.51$, $p < 0.01$. Due to the high correlation between response time and run time, and because run time was also the measure used by Capaldi and Miller (1988) to investigate the performance of rats in a similar study, we focused only on this measure.

The monkeys were not restrained in any way during this task, and occasional disengagement from the task in the middle of a trial resulted in unrealistically long start or response times. This caused the mean times to be much greater than the medians. We recorded all trials with run times in excess of 10 s to be false trials because this length was about three times the length of the typical trial. To ensure that the exclusion of these trials was justified we analyzed the medians for start time, response time, and total time. The medians were comparable to the means obtained when excluding these trials, so the 10 s trial limit was used in all subsequent analyses.

With the exclusion of these trials, Lou's run times during training ranged from 3.98 to 9.98 s with mean ± SD = 5.54 ± 1.01 s on reinforced trials and 4.20 to 7.87 s with mean ± SD = 5.56 ± 0.80 s on nonreinforced trials. Murph's run times during training ranged from 4.16 to 9.94 s with mean ± SD = 5.88 ± 0.89 s on reinforced trials and 4.70 to 9.57 s with mean ± SD = 6.12 ± 0.87 s on nonreinforced trials.

After two sessions of training, the data from both monkeys were analyzed to determine whether they should be moved to the test phase of the experiment. Both monkeys performed significantly slower on the fourth (nonreinforced) trials compared to the third (reinforced) trials during training [Murph: $F(1, 576) = 14.49$, $p < 0.01$, $\eta^2 = 0.03$; Lou: $F(1, 429) = 25.05$, $p<0.01$, $\eta^2 = 0.06$] so they were moved to test trials.

The training trials for both monkeys then were divided into blocks of 100 trials to assess progress over time. Figure 4.2 shows that by block 2, Lou developed a pattern of running slowly on the first trial: mean ± SD = 5.77 ± 0.91 s, faster on the next two trials: mean ± SD = 5.61 ± 1.14 s, 5.06 ± 0.67 s, and slower on the last nonreinforced trial: mean ± SD = 5.49 ± 0.62 s. This pattern persisted through the next block of trials as well.

By block 2, Murph developed a pattern of performing slowly on the first two trials: mean ± SD = 5.97 ± 0.76 s, 5.96 ± 0.97 s, faster on the third trial: mean ± SD = 5.82 ± 0.86 s, and slower again on the last nonreinforced trial: mean ± SD = 6.15 ± 1.01 s, which persisted through block 3. During block 4, however, Murph developed a pattern of performing progressively slower on trials 2, 3, and 4: mean ± SD = 5.49 ± 0.48 s, 5.98 ± 0.79 s, 6.08 ± 0.63 s, of each series.

Lou's run times during testing ranged from 3.78 to 9.99 s with mean ± SD = 5.45 ± 0.9 s on reinforced trials and 4.36 to 9.93 s with mean ± SD = 5.71 ± 0.98 s on nonreinforced trials.

Murph's run times during testing ranged from 4.16 to 9.92 s with mean ± SD = 5.83 ± 0.87 s on reinforced trials and 4.87 to 9.73 s with mean ± SD = 6.19 ± 0.76 s on nonreinforced trials.

The first 10 probe trials of each novel numeral provided some evidence that immediate generalization to the new numerals occurred (see Table 4.1). For both monkeys, the average run time of the first 10 probe trials (excluding trials exceeding 10 s) is greater for the last nonreinforced trial in each novel series than the reinforced trial before it.

Figure 4.3 shows the mean run time for each trial number in each type of series for all 1,500 test trials. For both monkeys, a one-way ANOVA (run time × trial number) was performed separately for each of the three target series (2, 3, and 4). Results for Lou revealed a significant difference in mean run times based on trial number for all three target numerals [target 2: $F(2, 362) = 12.85$, $p < 0.01$, $\eta^2 = 0.07$; target 3: $F(3, 513) = 12.85$, $p < 0.01$, $\eta^2 = 0.07$; target 4: $F(4, 363) = 6.38$, $p < 0.01$, $\eta^2 = 0.07$]. Tukey post-hoc tests revealed that Lou's performance on the last (nonreinforced) trial in each series was significantly slower ($p < 0.05$) than the next to last (reinforced) trial for target numerals 2 and 3. On series involving the target numeral 4, Lou's performance was significantly faster on trial numbers 2, 3, and 4 compared to trial number 1. He was also significantly faster on trial number 3 compared to trial numbers 1 and 4.

The data for Murph also showed a significant difference in mean run times based on trial number for all three target numerals [target 2: $F(2, 218) = 7.23$, $p < 0.01$, $\eta^2 = 0.06$; target 3: $F(3, 464) = 21.43$, $p < 0.01$, $\eta^2 = 0.12$; target 4: $F(4, 561) = 13.63$, $p < 0.01$, $\eta^2 = 0.09$]. Tukey post-hoc tests revealed that Murph's performance on the last (nonreinforced) trial in each target series was significantly slower ($p < 0.05$) than his performance on the next to last (reinforced) trial in each series.

To investigate whether the monkeys were using temporal cues to predict the nonreinforced trial, a correlation was performed using run time on nonreinforced trials and run time on the previous reinforced trials for both monkeys. These variables were not significantly correlated for either monkey [Murph: $r(689) = -0.01$, $p = 0.71$; Lou: $r(639) = 0.07$, $p = 0.10$].

Discussion

With training, both monkeys developed a pattern of performing more slowly on the nonreinforced trial in an RRRN series compared to the reinforced trial before it. Lou developed the slow, fast, fast, slow pattern that Capaldi and Miller (1988) and Burns et al. (1995) observed in rats after training with an RRRN series. This provides evidence that Lou was anticipating the nonreinforced trial.

Murph developed a pattern similar to Lou's during blocks 2 and 3, but during block 4 he performed progressively slower on trials 2, 3, and 4 of the series. Although this caused him to perform significantly slower on the last (nonreinforced) trials compared to the next to last (reinforced) trials, the pattern does not provide evidence that he was predicting exactly when the nonreinforced trial would occur.

After being trained on a target 3 series, the monkeys showed signs of generalization to new target numerals 2 and 4. During the first 10 probe trials of each novel series, both monkeys averaged a slower run time for the last nonreinforced trial in the series than the reinforced trial before it. This was not accomplished by performing gradually slower on each trial in the series. This indicates that at the start of testing the monkeys may have already understood the importance of the Arabic numerals as a cue to the number of reinforced trials in each series. It also indicates that the monkeys were applying previously acquired knowledge of Arabic numerals to a novel task.

Throughout testing, Lou performed significantly slower on the nonreinforced trial compared to the reinforced trial before it for target series 2 and 3. He did not accomplish this by performing progressively slower on each trial in a series. For target series 4, however, his performance became gradually slower after trial number 3. Although Lou may not have been anticipating precisely when the nonreinforced trial would occur in the target 4 series, his gradually increasing time indicates that he may have been predicting a nonreinforced trial to occur at some point after trial 3.

Murph, on the other hand, learned to perform significantly slower on the nonreinforced trial compared to the reinforced trial before it for all three target series (2, 3, and 4). Despite the pattern developed in training, Murph did not perform progressively slower on each trial in those series. Instead, his run time after the first trial remained relatively stable until increasing significantly for the nonreinforced trial.

Lou's performance on target series 2 and 3, and Murph's performance on all three target series indicate that both monkeys were using the changing target numerals visible during testing to help them distinguish between series and predict when the nonreinforced trial would occur. Based on these encouraging findings and the evidence of generalization to new target numerals, we designed a second experiment to replicate these findings and to test the performance of monkeys presented with the full range of Arabic numerals on which they had been previously trained. We predicted that these new monkeys would show results similar to those Murph and Lou produced when presented with small target numerals in experiment 1. Due to the increased difficulty involved in keeping track of a greater number of trials, however, we were unsure whether the monkeys would have similar success with larger target series.

Experiment 2

Method

Subjects. Two male rhesus monkeys (*Macaca mulatta*), Willie and Gale, participated in this study. New monkeys were used in experiment 2 in an effort to replicate the results of experiment 1 with a separate group of animals. The new monkeys were 18 and 20 years old, respectively, and had testing histories similar to Murph and Lou in experiment 1. The monkeys were individually housed according to federal animal housing standards and were not food or water deprived during this study.

Apparatus, task, and training procedure. The apparatus, training task, and training procedure were identical to those used in experiment 1. After several sessions of training the data were analyzed to determine whether the monkeys should be moved to test phase or continue training using only the Arabic numeral 3. As was the case in experiment 1, the monkeys were considered ready for test phase if they showed significantly slower response times on the fourth (nonreinforced) trials compared to the third (reinforced) trials. Gale received 1,040 training trials (260 series) and Willie received 724 training trials (181 series) over the course of two sessions before they were moved to the test phase of the experiment.

Testing procedure. During testing, probe trials consisting of the Arabic numerals 2-8 were randomly interspersed with the familiar Arabic numeral 3 trials. A new numeral was randomly selected and introduced every 100 trials. As was true during training, the Arabic numeral displayed during testing corresponded to the number of reinforced trials that would occur before one nonreinforced trial. For example, the target numeral 5 corresponded to an RRRRRN series. In total, Willie completed 3,472 test trials over the course of seven sessions and Gale completed 6,853 test trials over the course of eight sessions.

Results

Data from experiment 2 were analyzed using the same procedure as experiment 1. Again, we excluded all trials with a run time in excess of 10 s. We analyzed the medians for each trial number to ensure that they were comparable to the means obtained when excluding the trials exceeding the 10 s limit. As was true in experiment 1, start time was somewhat correlated with run time, $r(12{,}089) = 0.41$, $p < 0.01$. As was also true in experiment 1, response time was highly correlated with run time, $r(12{,}089) = 0.83$, $p < 0.01$, so only run time was used in subsequent analyses.

Gale's run times during training ranged from 2.42 to 6.43 s with mean ± SD = 3.51 ± 0.69 s on reinforced trials and 2.68 to 9.94 s with mean ± SD = 3.83 ± 0.83 s on nonreinforced trials. Willie's run times during training ranged from 2.68 to 9.94 s with mean ± SD = 3.84 ± 0.99 s on reinforced trials and 2.59 to 9.27 s with mean ± SD = 3.96 ± 1.20 s on nonreinforced trials.

After two sessions of training, the data from both monkeys were analyzed to determine whether they should be moved to the test phase of the experiment. Gale performed significantly slower on the fourth (nonreinforced) trials compared to the third (reinforced) trials during training, $F(1, 479) = 4.05$, $p < 0.05$, $\eta^2 = 0.01$, so he was moved to test trials. Willie did not perform significantly slower on the fourth trials compared to the third trials when all of the training trials were combined; rather he performed significantly slower on the fourth trials compared to the third trials during the second block of 100 training trials, $F(1, 162) = 5.66$, $p < 0.05$, $\eta^2 = 0.03$. During the second block of 100 trials he also developed a pattern of performing slowly on the first trial: mean ± SD = 3.71 ± 0.91 s, faster on the next two trials: mean ± SD = 3.66 ± 0.98 s, 3.55 ± 0.74 s, and slower on the last nonreinforced trial: mean ± SD = 3.97 ± 1.43

s, indicating that he was anticipating when the nonreinforced trials would occur. Based on the statistical significance, as well as the general pattern of Willie's data from the second block of trials, Willie was moved to test trials as well.

Figure 4.4 shows the training trials for both monkeys, divided into blocks of 100 trials. Unlike Willie, Gale developed a pattern of performing progressively slower on trials 2, 3, and 4: mean ± SD = 3.18 ± 0.46 s, 3.47 ± 0.76 s, 3.93 ± 0.87 s, during the third block of training trials.

Gale's run times during testing ranged from 2.43 to 5.99 s with mean ± SD = 3.52 ± 0.54 s on reinforced trials and 2.52 to 5.98 s with mean ± SD = 3.64 ± 0.58 s on nonreinforced trials. Willie's run times during testing ranged from 2.45 to 7.42 s with mean ± SD = 3.91 ± 0.89 s on reinforced trials and 2.57 to 7.42 s with mean ± SD = 3.98 ± 0.92 s on nonreinforced trials.

The average run times for the first 10 probe trials of each novel numeral provided little evidence that the monkeys generalized to the new numerals. The monkeys performed more slowly on the nonreinforced trial for some novel series, but this was not a consistent pattern.

Figure 4.5 shows the mean run times for both monkeys on each type of testing series. During testing, Gale showed the same pattern of steadily increasing time for each trial in the target 3 series as he did in training. A similar pattern can also be seen for the target 2 and target 4 series. Gale did not develop a recognizable pattern on series with target numbers 5-8. One-way ANOVAs (trial number × run time) for each target series revealed a significant difference in mean run time based on trial number for target series 3 and 6 [target 3: $F(3, 946) = 5.05$, $p < 0.01$, $\eta^2 = 0.02$; target 6: $F(6, 996) = 2.38$, $p < 0.05$, $\eta^2 = 0.01$]. Tukey post-hoc tests showed that for target series 3, Gale's mean run times for trials 1 and 2 were significantly slower compared to trial 4 and for target series 6, Gale's mean run time was significantly slower for trial 1 compared to trial 6 ($p < 0.05$).

Willie showed the same pattern in the target 3 series during testing as he did during training with run times for trials 1 and 4 being the longest. The peak at trial 4 is present in every other target series except 7. One-way ANOVAs (trial number × run time) for each target series revealed a significant difference in mean run time based on trial number for target series 8, $F(8, 934) = 2.31$, $p < 0.05$, $\eta^2 = 0.02$. A Tukey post-hoc test for target series 8 showed that Willie's performance was significantly slower on trial 3 compared to trials 7 and 9 ($p < 0.05$).

One-way ANOVAs (target series × run time) for both monkeys revealed significant differences in mean run times based on the target series [Gale: $F(3, 946) = 5.05$, $p < 0.01$, $\eta^2 = 0.05$; Willie: $F(6, 5088) = 24.91$, $p < 0.01$, $\eta^2 = 0.03$]. Tukey post-hoc tests revealed that, in general, mean run times for both monkeys were faster on series with higher target numbers (see Table 4.2).

For both monkeys, a correlation was performed using run time on nonreinforced trials and run time on the previous reinforced trials to investigate the possibility that they were using temporal cues to predict the nonreinforced trial. These variables were not significantly correlated for either monkey [Gale: $r(679) = -0.03$, $p = 0.52$; Willie: $r(185) = 0.01$, $p = 0.93$].

Discussion

Like Murph in experiment 1, Gale developed a pattern of performing progressively slower on each trial in the RRRN series during training. Although this caused his mean times for nonreinforced trials to be higher than his mean times for reinforced trials, it does not provide evidence that he was predicting precisely when the nonreinforced trial would occur.

Willie's slow, fast, faster, slow pattern of performance during training was very similar to the pattern developed by Lou in experiment 1. This pattern indicates that he was predicting when the nonreinforced trial would occur. The established pattern also can be seen at the beginning of some of the longer series presented during testing. This indicates that Willie continued to use the strategy learned during training, even when it was no longer appropriate.

We anticipated that for some of the target series presented during testing the monkeys would use the changing target numerals (2-8) to develop a strategy of performing quickly on the reinforced trials and slowly on the nonreinforced trial. This was the strategy adopted by both Murph and Lou in experiment 1. Instead, the monkeys performed more slowly overall on series with higher target numbers. One possible explanation for this behavior is that the monkeys recognized that during series with higher target numbers, more reinforced trials occurred before the one nonreinforced trial. This might have motivated the monkeys to perform faster overall on series with higher target numbers.

General Discussion

During training, all four monkeys developed a pattern of performing more slowly on the nonreinforced trial in an RRRN series than on the reinforced trial before it. The patterns developed by the monkeys were different, however, and may indicate different strategies for performing this task. Lou and Willie developed the slow, fast, fast, slow pattern similar to that observed in rats trained on an RRRN series (Capaldi and Miller 1988; Burns et al. 1995). This provides evidence that these two monkeys were anticipating the nonreinforced trial. It is unclear why the monkeys ran slowly on the first trial in each series, but it is possible that they were slightly less motivated to perform the task after receiving no reward on the previous trial (the terminal N trial of the previous series). It is also possible that the monkeys took breaks during the inter-series intervals, which slightly delayed the start of each series.

Although two of the monkeys developed a pattern of responding similar to the pattern observed in rats (Capaldi and Miller 1988; Burns et al. 1995), both monkeys performed several hundred series before this pattern emerged. In contrast, the rats developed the pattern after performing less than 50 series. The monkeys may have required more training because of their extensive test histories. In most tasks previously performed by the monkeys a nonreinforced trial signaled an incorrect response. Therefore, the monkeys had to overcome the prior meaning of a nonreinforced trial before learning to predict when it would occur. The disparity in number of series required to develop the pattern may also indicate that the monkeys and rats were using different processes to perform this task.

There are several possible cues Lou and Willie may have been using to predict when the nonreinforced trial would occur. For instance, the monkeys initiated the trials themselves and there was no strict control of temporal cues so it is possible that they were using the duration of the first three trials to predict when the fourth nonreinforced trial would occur. The run times for the reinforced trials varied, however, with standard deviations during testing and training ranging from 0.54 to 1.01 s for the four monkeys. Therefore, run time was an imprecise cue and errors would likely occur. Those errors would most likely cause the occasional misjudgment of trial 3 as the final trial and therefore manifest themselves as an increased average time for trial 3. Lou and Willie showed a much faster time on trial 3 than any other trial in the RRRN series so it is unlikely that they were using the duration of the reinforced trials to predict when the nonreinforced trial would occur. In addition, run time on the nonreinforced trial could not be predicted by the total amount of time on the reinforced trials before it for any of the monkeys.

The naturally occurring variation in trial times also argues against the rhythm method (Davis and Pérusse 1988) of anticipating the nonreinforced trial. It is possible, however, that a larger variation is needed to disrupt the formation of a rhythmic pattern cue. In future studies, the inter-trial intervals could be varied systematically to test this hypothesis more directly.

Although response effort accumulated over the nonreinforced trials, it was also an unreliable metric for predicting when the nonreinforced trial would occur. Given that the monkeys performed thousands of trials a day, we would expect natural variations in arousal, fatigue, and hunger. Individual maze trials also varied unpredictably in effort any time the monkeys brought the cursor into contact with a wall of the maze during the solution. These natural variations would cause variations in response effort for individual trials.

In light of the arguments against the use of temporal, rhythmic, and accumulated effort cues, one explanation for the performance pattern shown by Lou and Willie on the RRRN series is that they were using numerical cues to predict when the nonreinforced trial would occur. Capaldi and Miller (1988) labeled this behavior as counting because they believed the rats were assigning abstract tags to the individual reinforced trials in accord with the one-to-one correspondence, stable-order, order irrelevance, and abstraction principles set forth by Gelman and Gallistel (1978) as the hallmarks of true counting.

In the current study and the previously discussed studies involving rats, the number of trials covaried with the number of rewards received. If the animals were responding based on numerousness the salient stimulus may have been the number of trials or the number of food pellets received. In either case, however, the theoretical implications would be the same.

Unlike Lou and Willie, Gale showed a pattern of increasing time for each trial of the RRRN series during training. Murph, who showed a pattern similar to Lou and Willie during the first part of training, also developed this pattern of increasing time during his last block of training. It is possible that Willie and Murph were anticipating a nonreinforced trial, but failed to use numerical information to pinpoint exactly which trial would be nonreinforced.

It is also possible that once rewarded for the first trial, Murph and Gale lost some interest in the task and therefore performed more slowly on the next trial. It is possible that this decrease in interest became greater with each trial, causing the final trial in the series to be the longest. According to this logic, the nonreinforced trial would have again stimulated interest and caused them to perform faster on the first trial in the next series. This scenario is unlikely given that both monkeys performed thousands of trials a day of this and other tasks, indicating very high motivation overall.

The clear pattern of increasing time per trial in each series most likely indicates that Murph and Gale were anticipating a nonreinforced trial occurring at some point after the initial trial. The use of an imprecise cue, such as the sum of the inter-trial times, response times, or run times would be expected to produce a pattern of increasing times for each trial. By using a temporal cue such as this, Murph and Gale would be less likely to mistake the first trial for the last trial, more likely to mistake the second trial for the last, and even more likely to mistake the third trial for the last. Mistakenly identifying a trial as nonreinforced should cause an increase in the run time for that trial.

It is also possible that Murph and Gale were using numerical cues to predict the nonreinforced trial without enumerating each individual trial and keeping track of the exact magnitude or cardinal value. Knowing that the nonreinforced trial occurs after “a few” reinforced trials would also cause errors in estimation. These errors would be expected to create a pattern of results similar to those observed for Murph and Gale.

This does not mean, however, that Murph and Gale are incapable of using numerical cues to predict precisely when an event will occur. In fact, Murph’s strategy changed during testing and he was able to predict when the nonreinforced trial would occur for an RRN, RRRN, and

RRRRN series, with only an Arabic numeral cue to help him distinguish between series. A possible reason for the use of an imprecise strategy involves the reinforcement schedule. The reward did not increase based on accuracy, so there was little motivation to anticipate the nonreinforced trial exactly. In this case, reliance on nonnumerical cues would provide a decent estimation of when the nonreinforced trial would occur.

Despite different performance strategies during training, both Murph and Lou learned to use the changing target numeral during testing to predict when the nonreinforced trial would occur. In fact, there is some evidence that the monkeys generalized the information obtained during training to new target numerals within the first 10 probe trials. Results from the entire testing phase show that Lou learned to predict the nonreinforced trial for target series involving the numerals 2 and 3 and Murph learned to predict the nonreinforced trial for target numerals 2, 3 and 4. This indicates that both monkeys recognized the connection between the Arabic numerals and the variable maze series, and incorporated those numerical cues into their performance strategies.

Although Willie and Gale did not use the changing target numerals (2-8) to predict precisely when the nonreinforced trial would occur in each series, they did respond differentially to the changing target numerals. The fact that both monkeys ran faster overall on series with higher target numerals indicates that Willie and Gale also incorporated the numeral cues into their strategy for performing this task.

Although this study illustrates the numerical competence of rhesus monkeys, there is no direct evidence that the monkeys were enumerating the individual maze trials with abstract tags, which is necessary for true counting as defined by Gelman and Gallistel (1978). It is possible that the monkeys were using an object-file (Uller et al. 1999) or accumulator mechanism

(Dehaene 1997; Gallistel and Gelman 2000) to perform this task. Although the results from the testing phase of experiment 2 could be used as evidence in favor of the object-file model because the monkeys were unable to predict the nonreinforced trial on target series higher than 4, we believe they are a better fit with the accumulator model. If the monkeys were using an object-file mechanism to store each individual trial in a slot in working memory, they would have performed at random when the slots became full. In experiment 2, however, both Willie and Gale developed a pattern of performance involving all the target numerals. This indicates that they were representing, at least in approximate form, the numerical value of the target numbers 2-8.

It should be noted that one major procedural difference distinguishes the experiments reported here from previous studies in which rats have been trained to anticipate nonreinforced trials in a series of maze runs. Whereas the rats were trained with one (or very few) series each day, the monkeys performed hundreds of trials each day on the maze task. A series of RRRN trials (cued by the numeral 3 as the target) was just four trials in an incredibly long sequence of reinforced and nonreinforced trials. Thus, the monkeys had to keep track of how many consecutive trials had been reinforced, relative to the target number for the trial, and also reset this sum with each new numeral sequence. Indeed, it is amazing that the monkeys even cared which trial would be nonreinforced, as even these N trials had to be completed for the animal to get to the next sequence of reinforced trials. Doubtless, this procedure contributed to the amount of variability observed in these data, and it seems reasonable to suggest that the effects would have been even cleaner and clearer if the monkeys had been tested with just a few distinct sequences each day.

Notwithstanding, this study provides the first evidence that number-trained rhesus monkeys can use Arabic numerals as a cue to help them perform a task involving sequential presentations, also known as a "constructive" enumeration task (Xia et al. 2000; Beran and Rumbaugh 2001). However, the pattern established by Murph and Gale during training of performing gradually slower on each trial in a series, and the failure of Gale and Willie to generalize the pattern learned during training to new target numerals, highlights the need for a task that specifically addresses the monkeys' understanding of when a series is finished.

.

References

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, Rumbaugh DM (2001) "Constructive" enumeration by chimpanzees (*Pan troglodytes*) on a computerized task. Anim Cogn 4:81-89

Boysen ST, Berntson GG (1989) Numerical competence in a chimpanzee (*Pan troglodytes*). J Comp Psychol 103:23-31

Brannon EM, Terrace HS (1998) Ordering of the numerosities 1-9 by monkeys. Science 282:746-749

Brannon EM, Terrace HS (2000) Representation of the numerosities 1-9 by rhesus monkeys (*Macaca mulatta*). J Exp Psychol Anim Behav Proc 16:31-49

Burns RA, Criddle CR (2001) Retention of ordinal position information with limited and extended serial training. Psychol Rec 51:445-452

Burns RA, Dunkman JR, Detloff SL (1999) Ordinal position in the serial learning of rats. Anim Learn Behav 27:272-279

Burns RA, Goettl ME, Burt ST (1995) Numerical discriminations with arrhythmic serial presentations. Psychol Rec 45:95-104

Burns RA, Johnson KS, Harris BA, Kinney BA, Wright SE (2004) Functional cues for position learning effects in animals. Psychol Rec 54:233-254

Call J (2000) Estimating and operating on discrete quantities in orangutans (*Pongo pygmaeus*). J Comp Psychol 114:136-147

Capaldi EJ, Miller DJ (1988) Counting in rats: its functional significance and the independent cognitive processes that constitute it. J Exp Psychol Anim Behav Proc 14:3-17

Capaldi EJ, Miller DJ (2004) Serial learning in rats: a test of three hypotheses. Learn Motiv 35:71-81

Davis H, Bradford SA (1986) Counting behavior by rats in a simulated natural environment. Ethology 73:265-280

Davis H, Pérusse R (1988) Numerical competence in animals: definitional issues, current evidence and a new research agenda. Behav Brain Sci 11:561-615

Dehaene S (1997) The number sense. Oxford University Press, New York

Emmerton J (1998) Numerosity differences and effects of stimulus density on pigeons' discrimination performance. Anim Learn Behav 26:243-256

Gallistel CR, Gelman R (2000) Non-verbal numerical cognition: from reals to integers. Trends Cogn Sci 4:59-65

Gelman R, Gallistel CR (1978) The child's understanding of number. Harvard University Press, Cambridge

Gulledge JP (1999) Judgments of dot arrays and Arabic numerals by Rhesus macaques: evidence of concept of number. Master's thesis, Georgia State University, Atlanta

Hauser MD, Carey S, Hauser LB (2000) Spontaneous number representation in semi-free-ranging rhesus monkeys. Proc R Soc Lond B 267:829-833

Hauser MD, Tsao F, Garcia P, Spelke ES (2003) Evolutionary foundations of number: Spontaneous representation of numerical magnitudes by cotton-top tamarins. Proc R Soc Lond B 270:1441-1446

Mandler G, Shebo BJ (1982) Subitizing: an analysis of its component processes. J Exp Psychol Gen 111:1-22

Matsuzawa T (1985) Use of numbers by a chimpanzee. Nature 315:57-59

Olthof A, Iden CM, Roberts WA (1997) Judgments of ordinality and summation of number symbols by squirrel monkeys (*Saimiri sciureus*). J Exp Psychol Anim Behav Proc 23:325-339

Piazza M, Mechelli A, Butterworth B, Price CJ (2002) Are subitizing and counting implemented as separate or functionally overlapping processes? Neuroimage 15:435-446

Rumbaugh DM, Richardson WK, Washburn DA, Savage-Rumbaugh ES, Hopkins WD (1989) Rhesus monkeys (*Macaca mulatta*), video tasks, and implications for stimulus- response spatial contiguity. J Comp Psychol 103:32-38

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

Smith JD, Shields WE, Washburn DA (2003) The comparative psychology of uncertainty monitoring and metacognition. Behav Brain Sci 26:317-373

Thomas RK, Fowlkes D, Vickery JD (1980) Conceptual numerousness judgments by squirrel monkeys. Am J Psychol 93:247-257

Uller C, Carey S, Huntley-Fenner G, Klatt L (1999) What representation might underlie infant numerical knowledge? Cogn Dev 14:1-36

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

Washburn DA (1994) Stroop-like effects for monkeys and humans: processing speed or strength of association? Psychol Sci 5:375-379

Washburn DA, Gulledge JP (2002) A species difference in visuospatial working memory in human adults and rhesus monkeys. Int J Comp Psychol 15:288-302

Washburn DA, Rumbaugh DM (1991) Ordinal judgments of numerical symbols by macaques

(*Macaca mulatta*). Psychol Sci 2:190-193

Washburn DA, Rumbaugh DM (1997) If faster is smarter, why are we slower: a comparative perspective on intelligence and processing speed. Am Psychol 52:1147-1148

Xia L, Siemann M, Delius JD (2000) Matching of numerical symbols with number of responses by pigeons. Anim Cogn 3:35-43

Author Note

This research was supported by a grant (HD-38051) from the National Institute of Child Health and Human Development to the Language Research Center of Georgia State University. The authors would like to thank Michael Beran and Jonathan Gulledge for their advice and assistance in this research. All applicable federal, disciplinary, and institutional rules and regulations regarding animal care and use have been followed in the care and testing of the monkeys.

Table 4.1

Average Run Time (in Seconds) for the First 10 Probe Trials of Each Novel Numeral

	Lou		Murph	
	Target 2	Target 4	Target 2	Target 4
Trial 1	5.67	5.67	5.84	6.21
Trial 2	5.41	5.60	5.79	5.80
Trial 3	5.61	6.00	5.96	5.57
Trial 4		5.14		5.62
Trial 5		5.30		6.15

Table 4.2

Mean Run Times (in Seconds) for Each Target Series during Testing

	Target 2	Target 3	Target 4	Target 5	Target 6	Target 7	Target 8
Gale	3.80_{d}	3.74_{d}	3.65_{c}	3.54_{b}	3.48_{ab}	3.46_{a}	3.47_{ab}
Willie	4.20_{c}	4.10_{bc}	4.03_{b}	3.88_{a}	3.76_{a}	3.78_{a}	3.84_{a}

Note. Means in the same row that do not share subscripts differ at $p < 0.05$ in a Tukey post-hoc comparison

Figure 4.1. The display used during training in Experiments 1 and 2. The goal was to move the cursor in the lower right hand corner to make contact with the numeral 3.

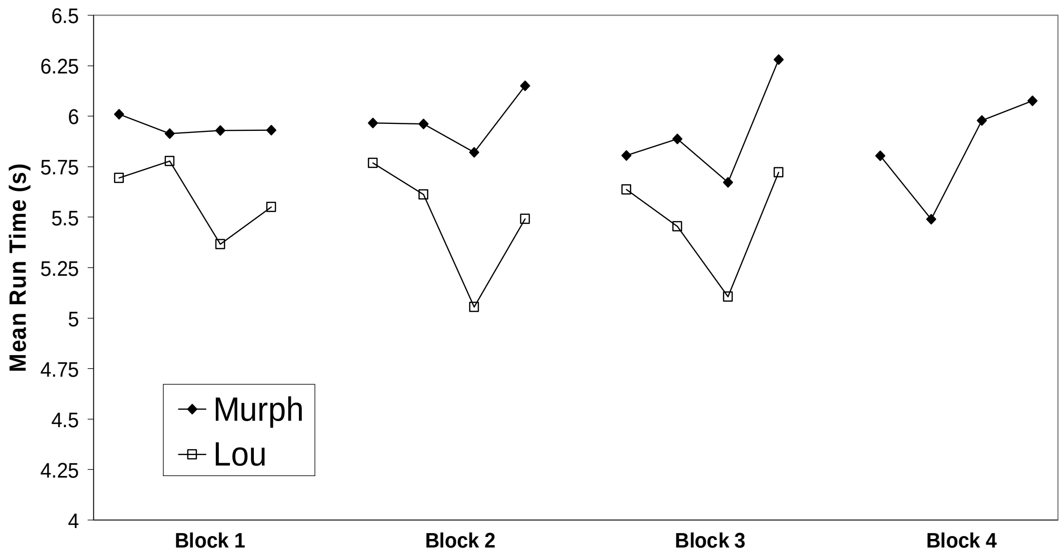


Figure 4.2. The mean run time for each trial number (in blocks of 100 trials) for the RRRN series used during training.

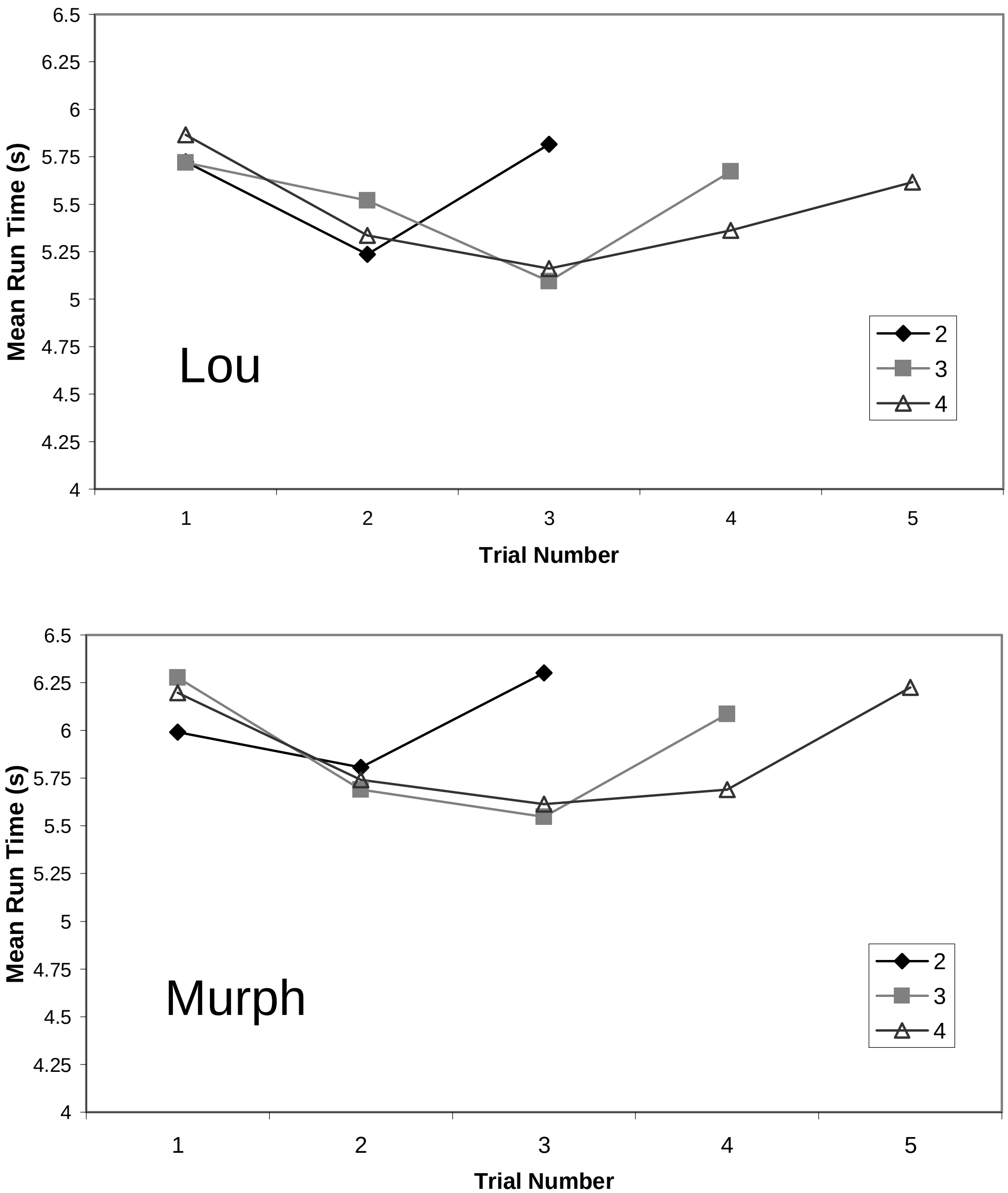


Figure 4.3. The mean run time for each trial number in each target series during testing. The numeral 2 in the legend denotes an RRN series, the numeral 3 an RRRN series, and the numeral 4 an RRRRN series.

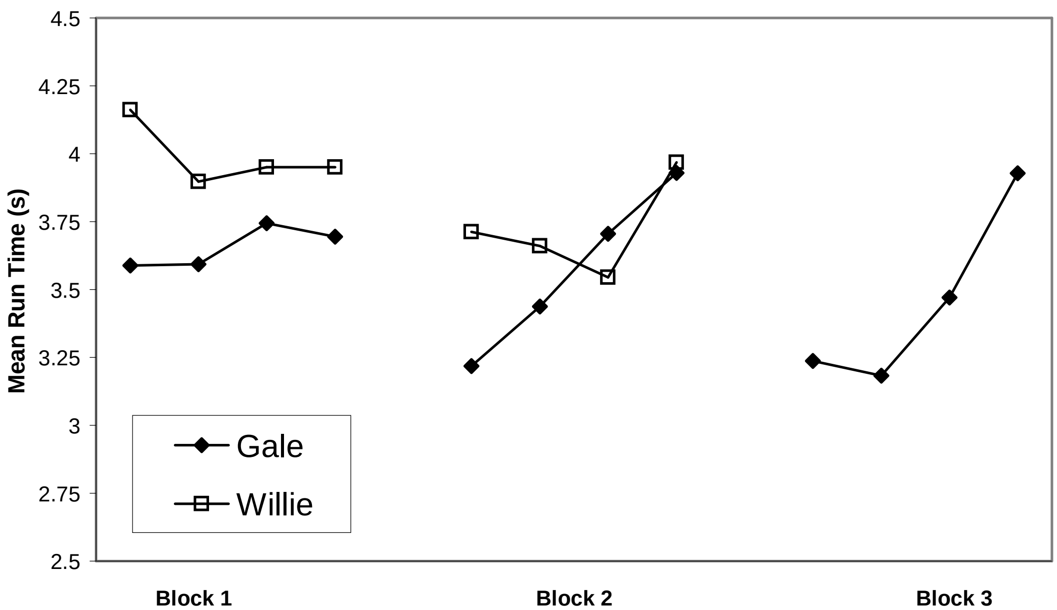
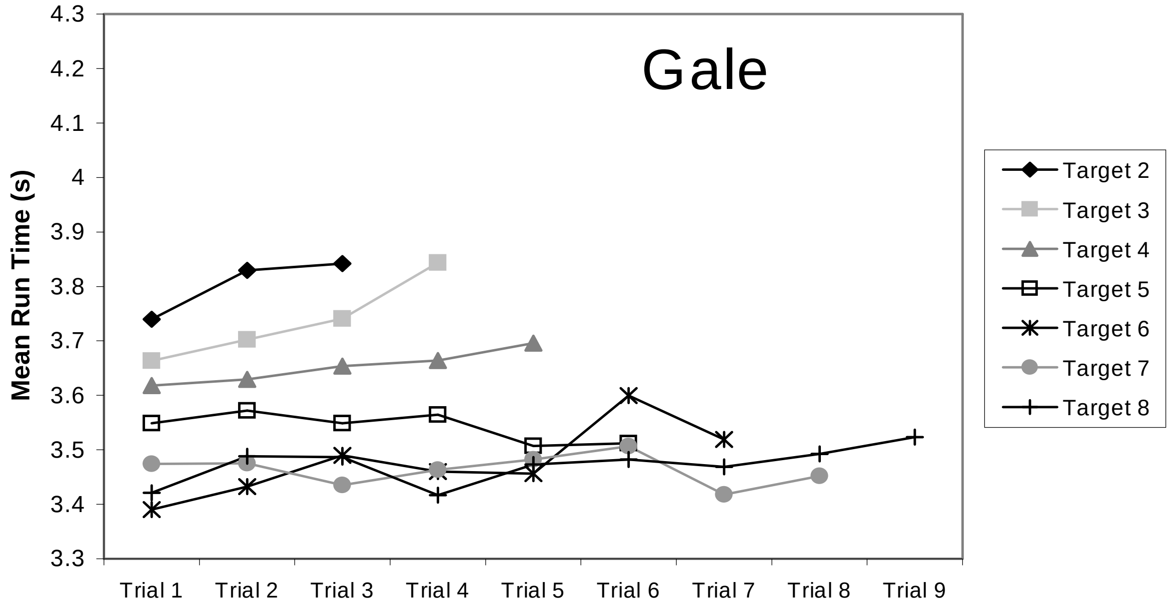


Figure 4.4. The mean run time for each trial number (in blocks of 100 trials) for the RRRN series used during training.



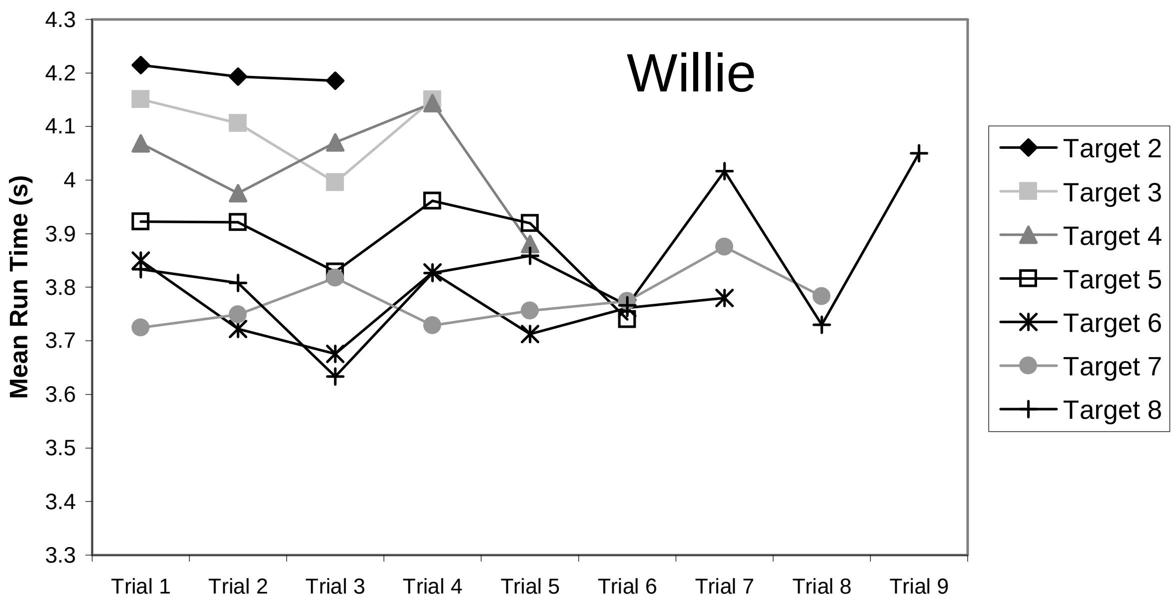


Figure 4.5. The mean run time for each trial number in each target series during testing.

Chapter 5: Rhesus monkeys (*Macaca mulatta*) select Arabic numerals or visual quantities corresponding to a number of sequentially completed maze trials[4]

Abstract

Four number-trained rhesus monkeys were trained to enumerate their sequential responses. After completing a series of computerized maze trials, the monkeys were given a same/different discrimination involving a numerical stimulus (an Arabic numeral or a visual quantity) and the letter *D*. The goal was to choose the numerical stimulus if it matched the number of just-completed maze trials, and to choose the letter *D* if it did not. There were large individual differences in performance, but one animal performed above 70% when receiving randomly intermixed series of 1, 3, 5, and 9 maze trials. This indicates that the monkey was keeping track of the approximate number of maze trials completed in each series and using that numerical cue to respond during the same/different discrimination.

[4] This chapter was previously published as: Harris, E.H., Washburn, D.A., Beran, M.J., & Sevcik, R.A. (2007). Rhesus monkeys (*Macaca mulatta*) select Arabic numerals or visual quantities corresponding to a number of sequentially completed maze trials. *Learning and Behavior, 35,* 53-59

Several different paradigms have been used to investigate numerical ability in animals. These include relative numerousness judgments, in which the animals choose between two or more sets of items on the basis of quantity (e.g., Beran, 2001; Boysen & Berntson, 1995; Brannon & Terrace, 2000; Call, 2000; Hauser, Carey, & Hauser, 2000; Nieder, Freedman, & Miller, 2002; Roberts & Mitchell, 1994; Thomas, Fowlkes, & Vickery, 1980) and tasks in which the absolute number of items is relevant (e.g., Beran & Rumbaugh, 2001; Boysen & Berntson, 1989; Capaldi & Miller, 1988; Davis, 1984; Emmerton, 1998; Matsuzawa, 1985; Murofushi, 1997; Pepperberg, 1994; Xia, Emmerton, Siemann, & Delius, 2001). The present study uses a paradigm in which the "to-be-enumerated" items are sequential events rather than visible items because we were interested in whether number-trained rhesus monkeys can match their own sequential responses with an Arabic numeral or visual dot quantity.

One of the many important aspects of human numerical competence involves the ability to keep track of sequentially presented items or events and to provide a numerical label corresponding to the cardinal value of the set. For example, adult humans asked to keep track of the number of traffic lights they pass on their way to work each morning would probably be able to provide the correct number. Several researchers have used the sequential presentation of items or events to investigate numerical ability and serial learning in rats (e.g., Burns & Criddle, 2001; Burns, Johnson, Harris, Kinney, & Wright, 2004; Capaldi & Miller, 2004). In one such study, Davis & Bradford (1986) trained rats to enter either the third or fourth tunnel in a series of six tunnels. The configuration of the tunnels and distance between them varied from trial to trial, so the only available cue was the number of previously encountered tunnels.

Capaldi and Miller (1988) trained rats with a three-trial series of maze runs consisting of two reinforced trials followed by a nonreinforced trial (RRN) and a four-trial (NRRN) series beginning and ending with a nonreinforced trial. The rats quickly developed a pattern of running more slowly on the terminal nonreinforced trial of each series than on the other, reinforced trials. This indicates that they were keeping track of the number of completed trials and predicting when the nonreinforced trial would occur. Burns, Goettl, and Burt (1995) systematically varied the intertrial intervals in a series of runway trials and concluded that the slower running times observed on the terminal nonreinforced trials could not be explained by rhythmic cues, as had been suggested by Davis and Pérusse (1988).

A recent study from our laboratory focused on the ability of four number-trained rhesus monkeys, including those in the present study, to use an Arabic numeral cue to predict when a nonreinforced event would occur (Harris & Washburn, 2005). The monkeys were presented with a computerized task consisting of three reinforced maze trials followed by one nonreinforced trial (RRRN). The goal of the maze was an Arabic numeral 3, which corresponded to the number of reinforced trials in the series. Two of the monkeys eventually developed a “slow, fast, faster, slow” pattern similar to that of the rats in the Capaldi and Miller (1988) study. Judging by the slow running time on the terminal nonreinforced trial of the series, the monkeys had been anticipating the nonreinforced trial. The other two monkeys performed gradually slower on each trial in the series, which made it difficult to speculate on their ability to predict the nonreinforced trial.

Two of the monkeys then were given probe series of the numerals 2 and 4, and the remaining two monkeys were given probe series of the numerals 2 through 8. These probe series were randomly intermixed with the familiar numeral 3 series. As was true during training, the Arabic numeral displayed in the maze corresponded to the number of reinforced trials that would occur before one nonreinforced trial. For instance, a numeral 2 indicated that it was an RRN series.

The monkeys receiving the probe series 2 and 4 showed generalization to the new numerals and developed a pattern of performing more slowly on the nonreinforced trial than on the reinforced trials before it, indicating the use of the changing target numeral to anticipate the nonreinforced trial. The monkeys receiving probe series of the numerals 2 through 8 did not use the changing numerals to predict precisely when the nonreinforced trial would occur in each series, but they did incorporate numerical cues into their performance strategy. They responded differentially to the targets by running faster overall on series with higher target numerals. One explanation for this result is that the monkeys recognized that a higher target numeral indicated more reinforced trials before the one nonreinforced trial. This may have motivated the monkeys to perform faster overall on those series.

Although not all of the monkeys in the Harris and Washburn (2005) study used the target numbers in the way anticipated, there was little motivation for the monkeys to keep track of the absolute number of trials. The reinforcement pattern remained the same, no matter what strategy the monkeys used to perform the task. During training, for instance, the monkeys always received three reinforced trials followed by one nonreinforced trial, regardless of how quickly they completed each maze trial. In addition, the monkeys performed thousands of trials a day on this task and other tasks, so a few nonreinforced trials were probably not very salient.

In the present study, the monkeys were required to compare the number of maze trials they had just completed to two choice options, and they were reinforced only when they made a correct response. This would increase motivation to perform at high levels because of the time invested in each series of maze trials. The current study is unique in that it not only tests the ability of number-trained monkeys to keep track of sequential events, but also tests their ability to compare numerical labels for cardinal values to sequentially completed responses that must be enumerated. Although chimpanzees have demonstrated an ability to label a visible quantity of items with an Arabic numeral (e.g., Biro & Matsuzawa, 2001; Boysen & Berntson, 1989; Matsuzawa, 1985; Murofushi, 1997; Tomonaga & Matsuzawa, 2002), this ability in rhesus monkeys has never been demonstrated with simultaneously visible items or sequentially completed events.

In this experiment, the monkeys received series of 1, 2, 3, 5, or 9 computerized maze trials, followed by two response options. One option was a numerical stimulus (either an Arabic numeral or a dot array) that either matched or differed from the number of maze trials that had been completed. The other option was a letter *D*, which represented "different" from the number of maze trials in the syntax of the computer program. We were interested in whether the monkeys could learn to choose the numerical stimulus when it matched the number of just-completed maze trials, or to choose the *D* when the numerical option did not match the number of just-completed maze trials. Because these animals had previously been trained to use Arabic numerals in quantity judgment tasks (e.g., Washburn & Rumbaugh, 1991), we also wanted to investigate any potential differences in performance as a function of the form that the numerical response option took (as either a numeral or a dot quantity). Given the previous manner in which the monkeys used numerals, we predicted that sequentially enumerated sets might be more easily represented as visual quantities, and that performance might be higher when the numerical response option took the form of a visual dot quantity. However, if Arabic numerals represented abstract quantities for the monkeys, then perhaps those stimuli also could be used appropriately within this task.

Method

Subjects

Four male rhesus monkeys (*Macaca mulatta*; Willie, Gale, Lou, and Hank) participated in this study. Their ages were 18, 20, 10, and 18 years, respectively. The monkeys were housed individually at the Language Research Center of Georgia State University according to federal animal housing standards. They were not deprived of food or water during this study.

All of these monkeys had participated in previous studies that required them to make relative numerousness and ordinal judgments using Arabic numerals and visual dot displays (e.g., Gulledge, 1999; Washburn & Rumbaugh, 1991). All except Hank also had experience in performing series of maze trials (Harris & Washburn, 2005). In addition, all four monkeys had participated in computerized joystick tasks related to various areas of cognitive research (e.g., Smith, Shields, & Washburn, 2003; Washburn & Gulledge, 2002; Washburn & Rumbaugh, 1997), including same/different judgments similar to those used in the current task, but with nonnumerical stimuli.

Design and Procedure

The monkeys were tested in their home cages using the Language Research Center Computerized Test System (see Rumbaugh, Richardson, Washburn, Savage-Rumbaugh, & Hopkins, 1989, for a description), which consists of a joystick attached to a computer and color monitor. The monkeys moved the joystick to control the movement of the cursor on the screen. The computer program recorded the target number along with the duration of each maze trial, the choices presented, and the accuracy and response time for each stimulus choice.

The computerized display consisted of a black H-shaped maze on a white background (Figure 5.1). The goal stimulus in the maze was a green rectangle appearing in one of four corners of the maze. The computer program randomly selected the corner on each trial. The monkey initiated the start of each trial by moving the joystick. At the beginning of each trial, the cursor appeared in the middle of the maze, and the monkey was required to move the cursor through the maze to the goal in order to complete a trial successfully.

Each series consisted of 1, 2, 3, 5, or 9 maze trials. The monkeys started with series of 1 and 9 maze trials, and additional series were added as the monkeys reached an accuracy criterion (see below). Once contact was made with the square at the end of the maze, the cursor returned to the middle of the maze and a new trial began with a new, randomly selected goal location. This meant that the animals could not use learned motor sequences for different numbers of maze trials because of the high number of variations of placement of the goals during each series. Completion of individual trials was not reinforced.

The monkeys involved in this study had been trained previously to pick the stimulus displaying the largest numerosity from an array. In contrast, this study required the monkeys to choose only numerical stimuli that matched the number of maze trials in a series. To avoid the monkeys' bias toward picking larger numbers, a same/different judgment was used instead of a matching-to-sample procedure.

Upon completion of all the maze trials in a series, the maze disappeared and two different stimulus choices immediately appeared on the screen. One choice (the numerical choice) was an Arabic numeral or visual quantity display and the other was a letter *D*, for "different." The *D* was white on a black background and was sized approximately 3 cm × 3 cm. The Arabic numerals also were white on a black background, and all were approximately 3 cm × 3 cm. The visual quantity display used randomly chosen white polygons of different sizes (hereafter referred to as "dots") on a black background. These polygons were unlike the round dot stimuli previously used with the monkeys in other numerical tasks. There were 10 different polygons used in the study, varying in approximate size from .5 and .75 cm. Each polygon in the visual array was placed in a random location within an invisible 5 × 5 matrix. Both the Arabic numerals and the dots were presented within a white 5 cm × 5 cm square.

The numerical stimulus presented during the labeling phase always corresponded to a possible number of maze trials from that test session. For example, when the monkeys were receiving only 1 or 9 maze trials in a series, the stimuli 1 and 9 (presented as numerals or dot quantities) were the only numerical stimuli choices used in the labeling phase. Within this constraint, the value of the Arabic numeral or number of dots was selected randomly by the computer program. The computer program randomly assigned the numerical stimulus to appear on the left or right side of the screen with the constraint that no more than four consecutive series could have the numerical stimulus displayed on the same side. The *D* appeared on the opposite side of the screen from the numeral or dot stimulus (Figure 5.2).

The type of numerical stimulus (numeral or dot quantity) that was presented also varied randomly from series to series, with the constraint that no more than four consecutive series could have the same type of stimulus. This was to ensure that one type of stimulus was not presented much more often than another, which would have caused the monkeys to form a bias toward the particular numeral or visual display, on the basis of the number of trials received.

If the numeral or visual dot quantity displayed during the labeling phase matched the number of maze trials in that series, the goal for the monkey was to move the cursor from the middle of the screen and make contact with that numerical stimulus. If the numeral or number of dots did not match the number of maze trials in the series, the goal was to move the cursor and make contact with the letter *D*.

Before each labeling phase, the computer program randomly determined whether the correct choice would be "same" or "different." Therefore, the correct choice was the dot quantity or numeral for approximately half of the discriminations and the letter *D* for the other half. This ensured that the monkeys needed to use the changing number of maze trials in order to be reinforced at a greater-than-chance level.

Correct responses during the labeling phase were rewarded with sound feedback and the automatic delivery of 94-mg fruit-flavored pellets. The number of pellets delivered corresponded to the number of maze trials in that series. For instance, a correct response after a two-trial series was rewarded with two pellets. Incorrect responses resulted in a 15-sec time-out and a negative buzzing sound. After the monkeys had completed the labeling phase of a series, a new series of maze trials began. To prevent the monkeys from developing a bias toward the *D* or the numerical stimuli, incorrect series were repeated until the monkeys made the correct choices.

The monkeys had continuous access to the task for several hours a day, several days a week. At the start of testing, they were presented with randomly intermixed series consisting of 1 or 9 maze trials and discriminations involving only the numerical stimulus 1 or 9 and the *D*. (We started the maze trials with the extreme values of 1 and 9 to aid the monkeys in conceptually connecting the maze trials and the discriminations.) Trials were administered in 100-series blocks. One additional numerosity (5, 3, or 2, in that order) was added each time the monkeys reached a performance level of 70% or better over the three most recent blocks. These additional numerosities were randomly intermixed with the familiar numerosities.

The additional numerosities (5, 3, and 2) were all chosen to facilitate learning of the task. The numeral 5 was chosen to take advantage of the distance effect, which suggests that discriminations are easier when two numbers are farther apart. The numerals 2 and 3 were chosen to take advantage of the magnitude effect, which suggests that when distance is held constant, discriminations are easier with smaller numbers compared to larger ones (Moyer & Landauer, 1967).

Testing ended for each monkey when it failed to reach criterion after 30 blocks (3,000 series) with a given set of randomly intermixed numerosities. Additional numerosities were not used because none of the monkeys achieved the accuracy criterion with the numerosities 1, 2, 3, 5, and 9.

Analyses

As stated previously, when a monkey gave an incorrect response, the series was repeated until a correct response was given. The first response was included in analyses, and all correction series were excluded.

The monkeys were not restrained during this task, so they occasionally took a break to rest, eat, drink water, utilize another enrichment device, or engage in social behavior. This could result in unrealistically long trial times, and it caused the mean times to be much greater than the medians. All series in which a maze trial lasted longer than 10 sec were excluded from analyses, because this duration was about three times that of the typical maze trial. Across the four monkeys, this resulted in the exclusion of an average of 3.73% of the series. To ensure that the exclusion of these series was justified, the medians for the maze trials were analyzed. The medians were comparable to the means obtained when excluding these series, so the 10-sec trial limit was used in all subsequent analyses.

Results

Table 5.1 shows the number of blocks (100 series each) required by each monkey to reach the 70% accuracy criterion after each new numerosity was added. All of the monkeys achieved the accuracy criterion when presented with only the numerosities 1 and 9, three of the four monkeys reached criterion with the numerosities 1, 5, and 9; and one monkey reached criterion with the numerosities 1, 3, 5, and 9. However, none of the monkeys was able to reach criterion when presented with the numerosities 1, 2, 3, 5, and 9. The number of trials required for the monkeys to reach criterion with two numerosities was not a good predictor of how well they performed overall on this task. For example, Hank required more trials than any other monkey to reach criterion with two numerals, but he was the only monkey to achieve criterion with four numerosities.

Only data from the last 3,000 series performed by the monkeys were used in subsequent analyses. These trials were chosen because they contain the greatest range of numerosities for each monkey and therefore provide the greatest opportunity for analyses relevant to the experimental questions and hypotheses. Although the monkeys did not achieve the 70% accuracy criterion for these trials, they had achieved the accuracy criterion for the previous set of numbers. Therefore, at the start of the 3,000 series, the monkeys were already performing at greater-than-chance levels with every numerosity except the most recent addition.

To assess possible practice effects over the course of the last 3,000 series (30 blocks), correlation coefficients were computed for each monkey to determine whether accuracy increased or decreased as block number increased. No significant correlations ($p < .05$, two-tailed) were found for any of the monkeys (Willie, $r = .11$; Lou, $r = .06$; Gale, $r = .24$; Hank, $r = -.18$), indicating that the performance of these monkeys did not change significantly. Thus, these 3,000 series represent full, mature performance on the task.

Figure 5.3 shows the percentage accuracy for each number of maze trials and each stimulus type. Although Gale and Willie both show significantly higher accuracy ($p < .05$) when presented with numerical stimuli in the form of Arabic numerals for at least one number of maze trials, there is no consistent pattern to indicate a meaningful interaction. For example, Gale and Willie were not consistently more accurate on numeral trials than on dot quantity trials for series with low, high, or intermediate numbers of maze trials. Lou showed the opposite pattern of performance for series with one maze trial, which provides further evidence that there is no meaningful interaction between stimulus type and target number.

A two-way ANOVA of the effect of stimulus and trial type on accuracy was performed, using data from all four monkeys. *Stimulus type* refers to the form in which the numerosity was presented (numeral or dot quantity) and *trial type* refers to the correct response required for that trial (the numerical stimulus or the *D*). Although power was low, this analysis yielded no significant differences [stimulus type, $F(1,3) = 3.33$, $p = .17$, $\eta^2 = .53$, observed power = .25; trial type, $F(1,3) = 1.63$, $p = .29$, $\eta^2 = .35$, observed power = .15; stimulus type × trial type interaction, $F(1,3) = 1.20$, $p = .35$, $\eta^2 = .29$, observed power = .12]. Descriptive statistics also indicate that there was no bias toward one type of stimulus or trial type [dots with a numerical stimulus response *M(SD)* = 62.52% (8.50%); dots with a *D* response *M(SD)* = 63.03% (8.87%); numerals with a numerical stimulus response *M(SD)* = 72.55% (3.36%); numerals with a *D* response *M(SD)* = 62.77% (3.81%)]. Given this result of no difference in performance as a function of the form of the numerical stimulus presented at the labeling phase of the series, trial type and stimulus type were combined for all subsequent analyses, unless otherwise noted.

To test for a distance effect, the accuracy of all four monkeys was regressed on the numerical difference between the number of maze trials completed and the numerical stimulus that was presented during the labeling phase. The regression analysis revealed that accuracy was positively associated with the difference [$F(1,15) = 48.55$, $p < .05$, $R^2 = .76$]. It must be noted, however, that practice effects may have contributed to this correlation. The monkeys had the most practice with the numerosities 1 and 9, which are also the two numerosities farthest apart in distance. The magnitude of the numerosities involved could also have affected the correlation. To take into account the magnitude of the numerosities as well as their numerical distance, accuracy was regressed on the ratio of the smaller numerosity to the larger numerosity used in each series. For example, if a monkey completed one maze trial and was presented with the numeral 5 during the labeling phase, the ratio would equal 0.2. This regression analysis revealed that accuracy was significantly associated with ratio [$F(1,20) = 63.39$, $p < .01$, $R^2 = .76$]. This effect is illustrated in Figure 5.4.

To determine whether the monkeys were using the combined duration of the maze trials, instead of their numerical value to perform this task, data were analyzed from all of the trials on which the monkeys chose the numerical stimulus during the discrimination trial. We chose to look at this question post hoc rather than controlling for duration experimentally, because manipulating the duration of the trial would have caused other factors, such as rate of maze completion, to covary with the number of trials in a series. If the monkeys were using duration as a cue to this task, incorrect trials in which the monkeys chose a numerosity smaller than the number of maze trials performed should have occurred when the total duration of the maze trials was shorter than it usually was when they responded correctly for a given number of maze trials (i.e., incorrect maze trial duration < mean correct maze trial duration). In contrast, incorrect trials in which the monkeys chose a numerosity larger than the number of maze trials performed should have occurred when the total duration of the maze trials was larger than it usually was for correct trials (i.e., incorrect maze trial duration > mean correct trial duration).

For each monkey, a one-way ANOVA assessing the effects of trial type on the total duration of the maze trials was performed for each number of maze runs. The types of trial were categorized as those in which the monkey chose a numerical stimulus larger than the number of maze runs, those in which the monkey chose a numerical stimulus smaller than the number of maze runs, and those in which the monkey correctly chose the numerical stimulus. The only significant effect was found for Willie. For trials in which he performed only one maze run, his maze trial durations were significantly shorter on trials in which he chose a numerical stimulus larger than the number of maze trials compared with trials in which he correctly chose the numerical stimulus [$F(1,405) = 4.49$, $p < .05$, $\eta^2 = .01$]. This effect is opposite to what was predicted for a strategy involving duration as a cue to the correct response. Thus, none of the monkeys used differences in maze completion duration as the cue for which stimulus to select during the labeling phase.

Discussion

During the course of this study, the monkeys learned to label a series of sequentially completed maze trials with the corresponding Arabic numeral or visual dot quantity (or a *D* if the numerical option was not equal). All of the monkeys learned to match randomly intermixed series of 1 or 9 maze trials with the correct Arabic numeral or visual quantity when tested with a same/different discrimination. This provides evidence that the monkeys understood the task on some level and conceptually connected the maze series with the same/different discriminations.

This part of the task, however, could be performed by representing the number of maze trials simply as “few” and “many” (or “one and “many”), without representing the number of maze runs as a specific quantity. In fact, Willie seems to have used one of those strategies throughout this experiment. Willie achieved an accuracy level of 70% fairly quickly when presented with series of one and nine maze trials, but he did not achieve the accuracy criterion when the numerosity 5 was added. His pattern of errors revealed that after performing one maze run he almost never chose the numerosities 5 or 9, but after performing five maze runs he was more likely to choose 9 than 1, and after performing nine maze runs he was more likely to choose 5 than 1. This indicates confusion between the numerosities 5 and 9 that was not present for the numerosity 1.

Two of the monkeys participating in this study achieved accuracies greater than 70% for the numerosities 1, 5, and 9, within the first 500 presentations, but they did not reach criterion when the numerosity 3 was added to the experimental set. Their ability to perform the task with three numerosities indicates that their representation of the maze runs went beyond a simple representation of “one” and “many.”

The fourth monkey in the study, Hank, performed the task with the numerosities 1, 3, 5, and 9, but failed to achieve the 70% accuracy criterion after the numerosity 2 was added to the set. It is important to note that these numerosities were randomly intermixed, so Hank never received blocks of series containing only one numerosity. Rather, each new series of maze trials could consist of any of the numerosities in the set.

The monkeys were reinforced for correct choices with a number of pellets equal to the number of maze trials performed in the just-completed series as a motivation to complete the longer series. This did cause a slight high-number choice bias on pairs of trials in which the distance and magnitude of the two numerosities were the same. For instance, the monkeys were less accurate on trials in which they ran one maze trial and were presented with the numerosity 9 than they were on trials in which they ran nine maze trials and were presented with the numerosity 1 (an average accuracy of 79% in the former case and 95% in the latter). However, this bias did not prevent any of the monkeys from exceeding 70% accuracy for trials in which they performed 1 maze run and were presented with the tempting numerosity 9. The bias was even less pronounced for the numerosities 2, 3, and 5.

There is some evidence that the monkeys were using an approximate and variable representation of the number of maze runs to perform this task. Accuracy increased as a function of distance between the number of maze trials and the numerosity presented during the discrimination. Accuracy also decreased as a function of the ratio of the smaller numerosity to the larger numerosity used in each series, as predicted by Weber's law. Although a greater amount of practice with the numerosities 1 and 9 as compared to other numerosity pairs may have contributed to this correlation, a distance effect and adherence to Weber's law would occur if the monkeys' numerical representations were composed of inexact magnitudes. This is because inexact magnitudes would be more difficult to compare when the numerosities were close in distance and/or large in magnitude (Dehaene, 1997; Gallistel & Gelman, 2000).

The monkeys' error patterns were not related to the amount of time they spent on the maze trials in each series. The monkeys did not tend to choose numerosities that were higher than the correct choice after spending more time than usual on a particular series; therefore, they were not using duration alone as a cue to performing this task.

It is interesting that the monkeys performed equally well when the numerical stimulus was in Arabic numeral or visual dot quantity form. Although the visual quantities provide more inherent numerical information than the numerals, the monkeys have had a variety of testing experiences involving Arabic numerals. Their ability to match a series of maze trials to either a visual quantity or an Arabic numeral indicates flexibility in their performance strategy.

The ability of the monkeys to perform this task is impressive, due to the working memory demands and the absence of perceptual cues, such as surface area or density, to aid in the formation of their numerical representations. The monkeys were required to form a representation of a series of events, which lacked standard perceptual features, to update this representation throughout the series of maze trials, and to keep this representation in working memory while they chose the appropriate stimulus during the same/different discrimination.

Previous studies have found that nonhuman primates are capable of representing, combining, and comparing nonvisible, sequentially presented sets of items (e.g., Beran, 2001; Call, 2000; Hauser et al., 2000). This experiment provides strong evidence that monkeys can enumerate, albeit approximately, their own sequential responses and can match the number of responses with the corresponding Arabic numeral or visual quantity.

References

Beran, M. J. (2001). Summation and numerousness judgments of sequentially presented sets of items by chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology, 115,* 181-191.

Beran, M. J., & Rumbaugh, D. M. (2001). "Constructive" enumeration by chimpanzees (*Pan troglodytes*) on a computerized task. *Animal Cognition, 4,* 81-89.

Biro, D., & Matsuzawa, T. (2001). Use of numerical symbols by the chimpanzee (*Pan troglodytes*): Cardinals, ordinals, and the introduction of zero. *Animal Cognition, 4,* 193-199.

Boysen, S. T., & Berntson, G. G. (1989). Numerical competence in a chimpanzee (*Pan troglodytes*). *Journal of Comparative Psychology, 103,* 23-31.

Boysen, S. T., & Berntson, G. G. (1995). Responses to quantity: Perceptual versus cognitive mechanisms in chimpanzees (*Pan troglodytes*). *Journal of Experimental Psychology: Animal Behavior Processes, 21,* 82-86.

Brannon, E. M., & Terrace, H. S. (2000). Representation of the numerosities 1-9 by rhesus macaques (*Macaca mulatta*). *Journal of Experimental Psychology: Animal Behavior Processes, 26,* 31-49.

Burns, R. A., & Criddle, C. R. (2001). Retention of ordinal position information with limited and extended serial training. *Psychological Record, 51,* 445-452.

Burns, R. A., Goettl, M. E., & Burt, S. T. (1995). Numerical discrimination with arrhythmic serial presentations. *The Psychological Record, 45,* 95-104.

Burns, R. A., Johnson, K. S., Harris, B. A., Kinney, B. A., & Wright, S. E. (2004). Functional cues for position learning effects in animals. *Psychological Record, 54,* 233-254.

Call, J. (2000). Estimating and operating on discrete quantities in orangutans (*Pongo pygmaeus*). *Journal of Comparative Psychology, 114*, 136-147.

Capaldi, E. J., & Miller, D. J. (1988). Counting in rats: Its functional significance and the independent cognitive processes that constitute it. *Journal of Experimental Psychology: Animal Behavior Processes, 14,* 3-17.

Capaldi, E. J., & Miller, D. J. (2004). Serial learning in rats: A test of three hypotheses. *Learning and Motivation, 35,* 71-81.

Davis, H. (1984). Discrimination of the number three by a raccoon (*Procyon lotor*). *Animal Learning and Behavior, 12*, 409-413.

Davis, H., & Bradford, S.A. (1986). Counting behavior by rats in a simulated natural environment. *Ethology, 73,* 265-280.

Davis, H., & Pérusse, 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.

Emmerton, J. (1998). Numerosity differences and effects of stimulus density on pigeons' discrimination performance. *Animal Learning and Behavior, 26*, 243-256.

Gallistel, C. R., & Gelman, R. (2000). Non-verbal numerical cognition: From reals to integers. *Trends in Cognitive Sciences, 4,* 59-64.

Gulledge, J. P. (1999). *Judgments of dot arrays and Arabic numerals by Rhesus macaques: Evidence of concept of number*. Unpublished master's thesis, Georgia State University, Atlanta.

Harris, E. H., & Washburn, D. A. (2005). Macaques' (*Macaca mulatta*) use of numerical cues in maze trials. *Animal Cognition, 8,* 190-199.

Hauser, M. D., Carey, S., & Hauser, L. B. (2000). Spontaneous number representation in semi-free-ranging rhesus monkeys. *Proceedings of the Royal Society of London Series B-Biological Sciences, 267,* 829-833.

Matsuzawa, T. (1985). Use of numbers by a chimpanzee. *Nature, 315,* 57-59.

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgments of numerical inequality. *Nature, 215,* 1519-1520.

Murofushi, K. (1997). Numerical matching behavior by a chimpanzee (*Pan troglodytes*): Subitizing and analogue magnitude estimation. *Japanese Psychological Research, 39,* 140-153.

Nieder, A., Freedman, D. J., & Miller, E. K. (2002). Representation of the quantity of visual items in the primate prefrontal cortex. *Science, 297,* 1708-1711.

Pepperberg, I. M. (1994). Numerical competence in an African grey parrot (*Psittacus erithacus*). *Journal of Comparative Psychology, 108,* 36-44.

Roberts, W. A., & Mitchell, S. (1994). Can a pigeon simultaneously process temporal and numerical information? *Journal of Experimental Psychology: Animal Behavior Processes, 20*, 66-78.

Rumbaugh, D. M., Richardson, W. K., Washburn, D. A., Savage-Rumbaugh, E. S., & Hopkins, W. D. (1989). Rhesus monkeys (*Macaca mulatta*), video tasks, and implications for stimulus-response spatial contiguity. *Journal of Comparative Psychology, 103,* 32-38.

Smith, J. D., Shields, W. E., & Washburn, D. A. (2003). The comparative psychology of uncertainty monitoring and metacognition. *Behavioral and Brain Sciences, 26,* 317-373.

Thomas, R. K., Fowlkes, D., & Vickery, J. D. (1980). Conceptual numerousness judgments by squirrel monkeys. *American Journal of Psychology, 93,* 247-257.

Tomonaga, M., & Matsuzawa, T. (2002). Enumeration of briefly presented items by the chimpanzee (*Pan troglodytes*) and humans (*Homo sapiens*). *Animal Learning and Behavior, 30*, 143-157.

Washburn, D. A., & Gulledge, J. P. (2002). A species difference in visuospatial working memory in human adults and rhesus monkeys. *International Journal of Comparative Psychology, 15,* 288-302.

Washburn, D. A., & Rumbaugh, D. M. (1991). Ordinal judgments of numerical symbols by macaques (*Macaca mulatta*). *Psychological Science, 2,* 190-193.

Washburn, D. A., & Rumbaugh, D. M. (1997). If faster is smarter, why are we slower?: A comparative perspective on intelligence and processing speed. *American Psychologist, 52,* 1147-1148.

Xia, L., Emmerton, J., Siemann, M., & Delius, J. D. (2001). Pigeons (*Columba livia*) learn to link numerosities with symbols. *Journal of Comparative Psychology, 115,* 83-91.

Author Note

This research was supported by a Grant HD-38051 from the National Institute of Child Health and Human Development to the Language Research Center of Georgia State University, and by a Rumbaugh Fellowship to E.H.H. as partial fulfillment for the degree of Master of Arts. The authors thank Eric J. Vanman for his thoughts and comments during this project. All applicable federal, disciplinary, and institutional rules and regulations regarding animal care and use have been followed in the care and testing of the monkeys. Address all correspondence concerning this article to E.H. Harris, Language Research Center, Georgia State University, 3401 Panthersville Road, Decatur, GA 30034 (e-mail: eharris11@gsu.edu).

Table 5.1

The Number of Blocks (100 Series per Block) Required for Each Monkey to Reach the 70% Accuracy Criterion on Each Set of Numerosities

	1,9	1,5,9	1,3,5,9
Willie	10		
Gale	3	3	
Lou	4	5	
Hank	13	4	10

Note. Empty cells indicate that the monkey did not reach criterion for that set of numerosities

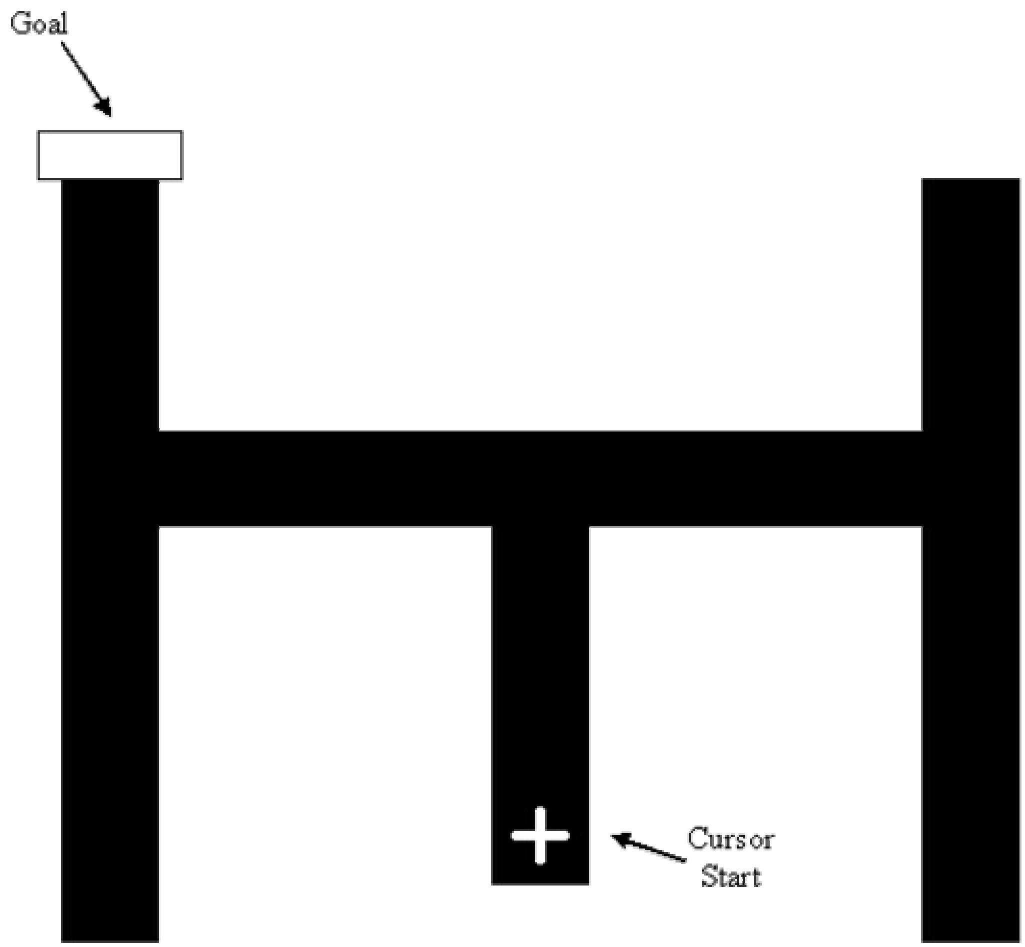


Figure 5.1. The maze display used during the series. The goal was colored green.

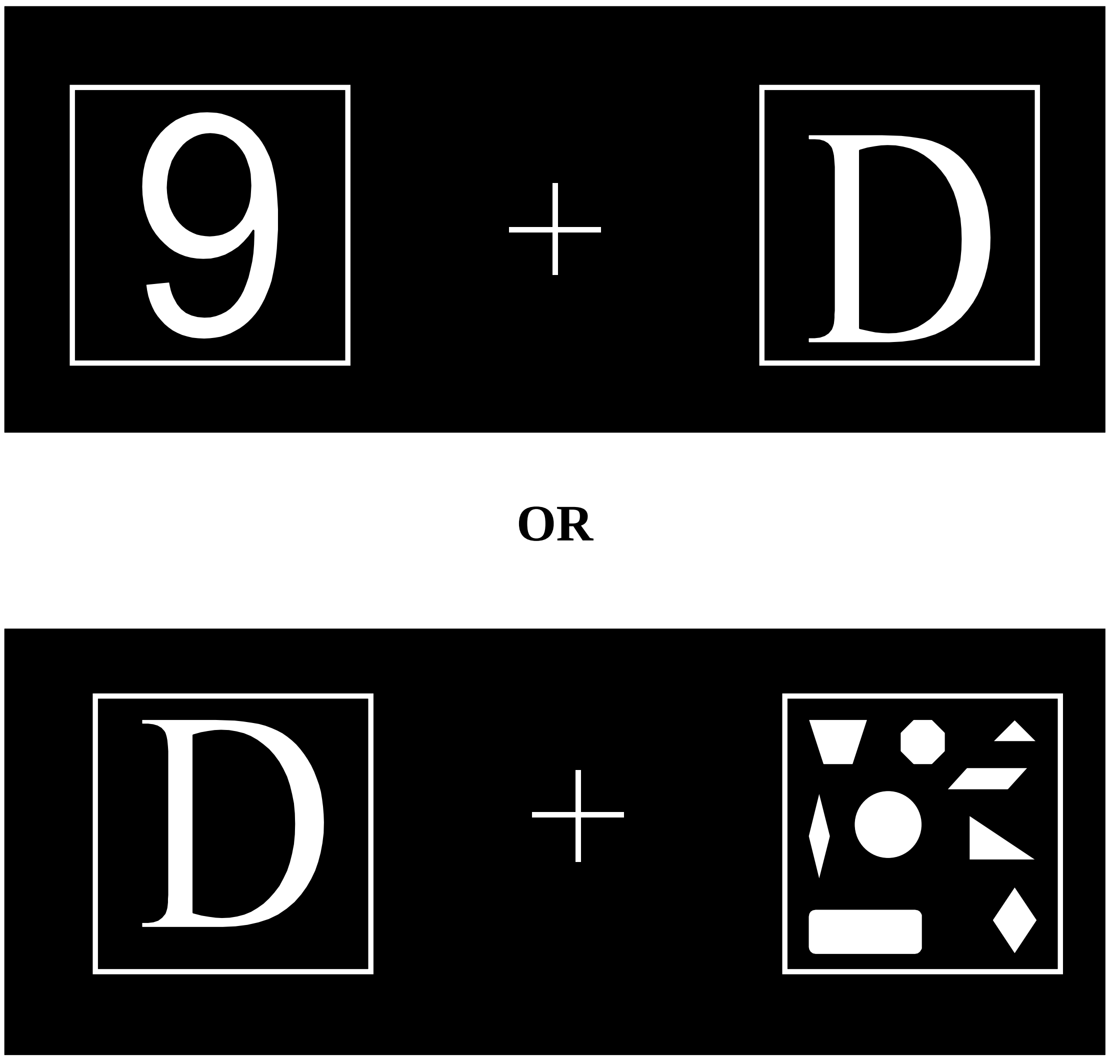
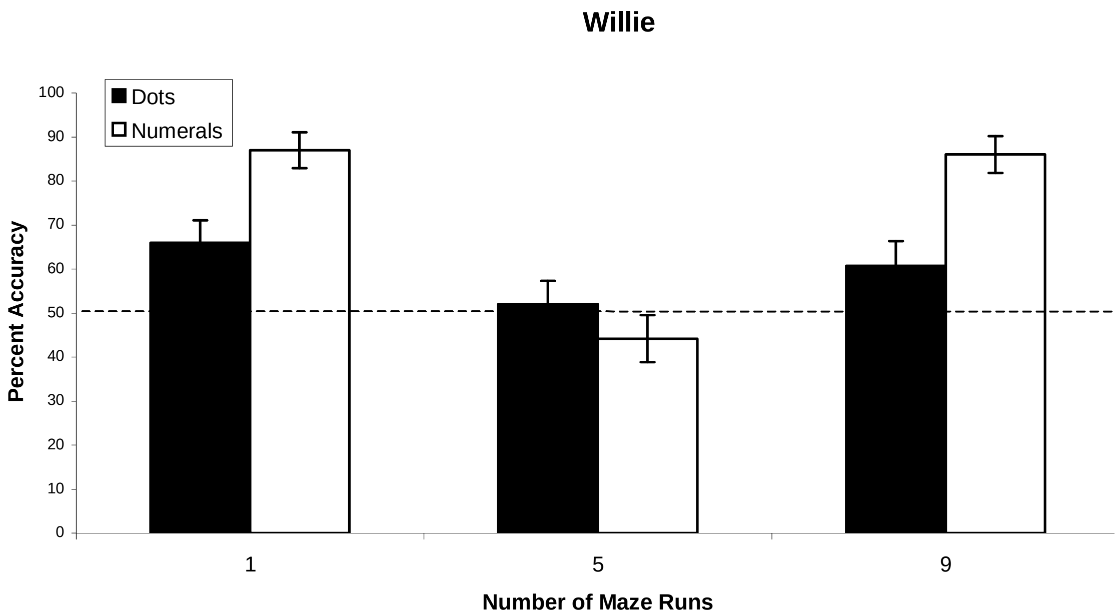


Figure 5.2. Example of display used during the labeling phase of a series. The + in the center of each figure is the cursor.



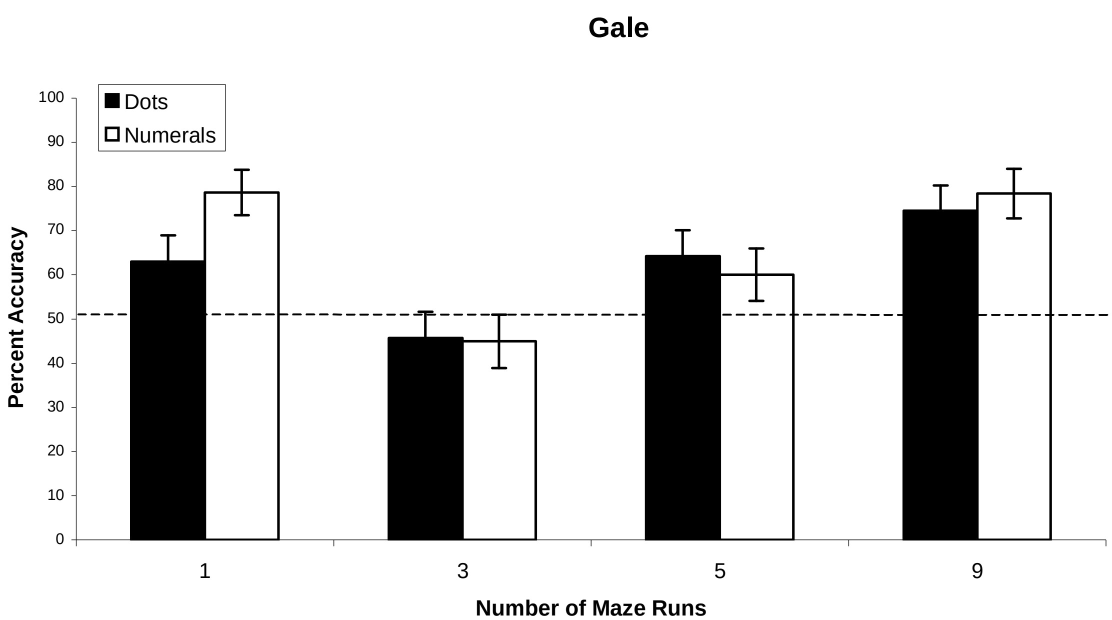


Figure 5.3. Percent accuracy for each type of stimulus and each number of maze runs. The dotted line denotes chance level. Error bars denote 95% confidence intervals.

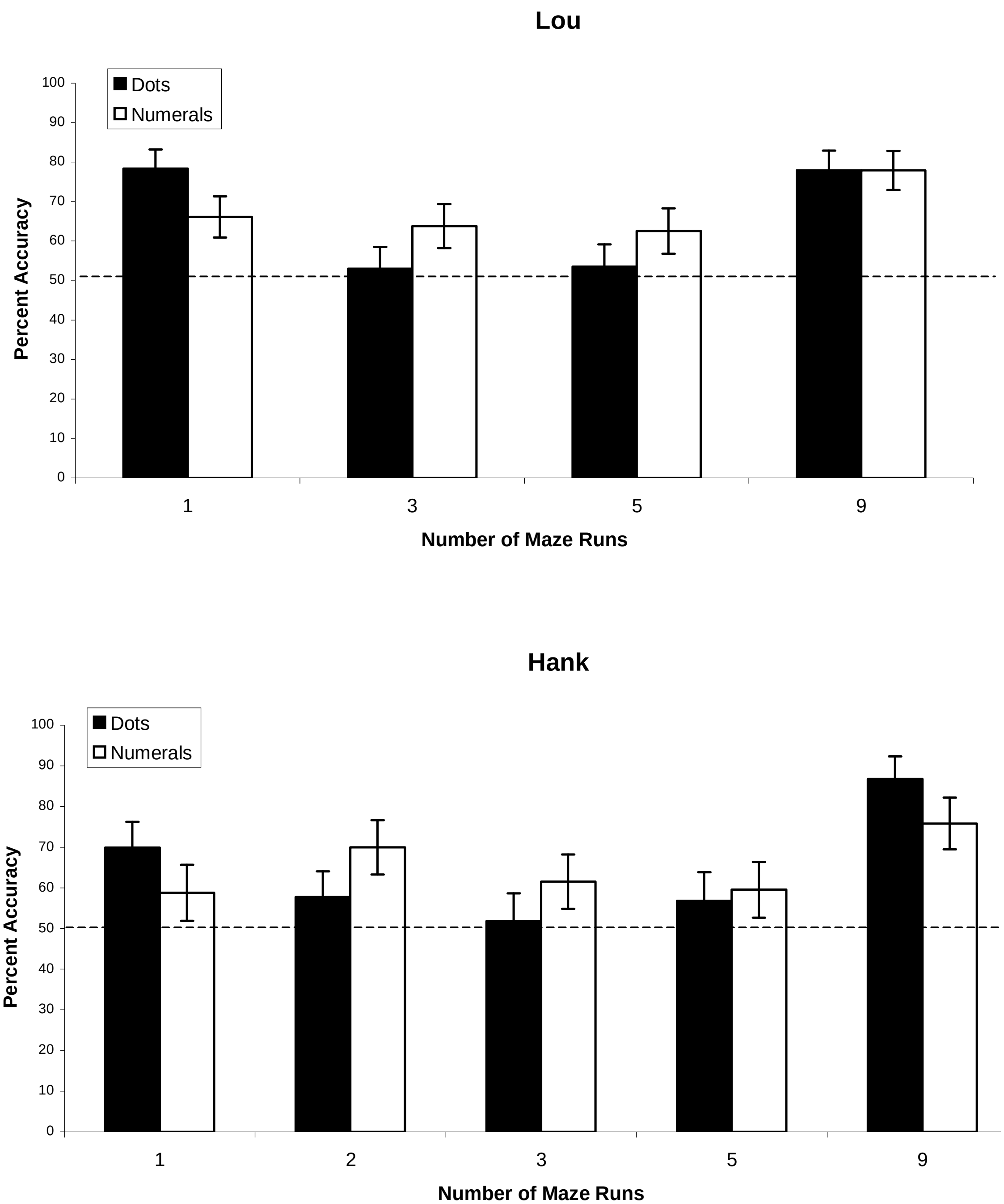


Figure 5.3. Continued

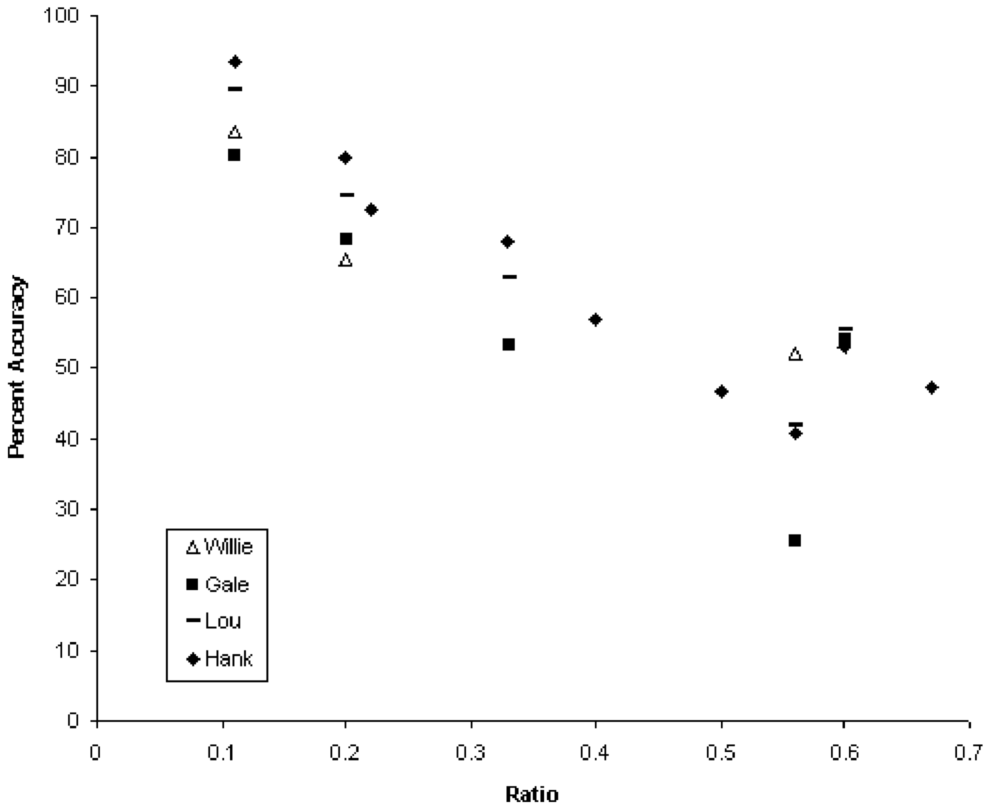


Figure 5.4. Accuracy as a function of the ratio of the smaller numerosity to the larger numerosity.

Chapter 6: Numerical abstraction across presentation mode by rhesus monkeys

Abstract

The ability of rhesus monkeys to transfer numerical rules learned in a sequential task to a simultaneous task was tested. In Experiment 1, eight monkeys were trained to make one response after viewing three sequentially presented circles on a computer screen and another response after seven sequential circles. During testing, the monkeys received nonreinforced simultaneous probe trials. Half of the monkeys showed some tendency to make a spontaneous "three" response after viewing three simultaneous circles and a "seven" response after viewing seven simultaneous circles, but only one monkey performed consistently above chance on the simultaneous trials. In Experiment 2, a different transfer paradigm was employed to investigate further the possibility of numerical transfer from a sequential to a simultaneous task, but no evidence of transfer was found. Overall, these experiments indicate that some monkeys can abstract number across different presentation modes, but this ability is exhibited only under limited conditions.

For adult humans, number is a broad category that can include objects, actions, and events that differ in perceptual features and modality. For example, three apples, three flashes of lightning, and three trips to the grocery store share the abstract numerical property of three. Gelman and Gallistel (1978) recognized abstractness as one of the five principles of formal counting and Davis and Pérusse (1988) argued that the ability to abstract number across different contexts and modalities is necessary for a true concept of number. Although adults routinely abstract number across different conditions, it is unclear whether this ability is shared by nonverbal populations with less numerical experience, such as human infants and nonhuman animals.

To date, only a small number of studies have been conducted to investigate numerical abstractness in infants and animals. In studies by Starkey and colleagues (Starkey, Spelke, & Gelman, 1983, 1990) infants were able to detect numerical correspondences between the visual and auditory modalities. While listening to a temporal sequence of two or three drumbeats, 6- to 9-month-old infants were presented with side-by-side photos depicting two and three household objects. Researchers found that when the infants were listening to three drumbeats they looked reliably longer at the visual display of three objects and when they were listening to two drumbeats they looked reliably longer at the display of two objects. Results were similar even when the duration of the two and three beat sequences were equated. In similar studies, Jordan and colleagues (Jordan, Brannon, & Gallistel, 2006; Jordan, Brannon, Logothetis, & Ghazanfar, 2005) provided evidence that rhesus monkeys and 7-month-old infants preferred to look at videos containing a number of conspecifics equal to the number of vocalizations they heard. These studies suggest that rhesus monkeys and very young human children possess an abstract concept of number that reaches across two sensory modalities.

Fernandes and Church (1982) presented rats with sequences of white noise and rewarded them for pressing the lever on the right when they heard two bursts of noise and the lever on the left when they heard four bursts of noise. Temporal cues were controlled by varying the duration of each burst as well as the total duration of the auditory sequences. With training the rats learned to respond based on the number of noise bursts. When the experimenters substituted light flashes for sounds, the rats immediately transferred their knowledge to the new task, which suggests that their representation of number was not tied to the auditory modality.

In a similar study, Church and Meck (1984) taught rats to press a lever on the left after viewing a sequence of two light flashes or hearing a sequence of two white noise bursts and a lever on the right after viewing a sequence of four light flashes or hearing a sequence of four noise bursts. When the rats were then presented with a combination of two lights and two noise bursts they spontaneously integrated the number of visual and auditory stimuli and responded by pressing the right lever. This indicates that the rats based their behavior on an abstract, amodal representation of number.

Davis and Albert (1987) trained rats on a more complex task that required them to discriminate between two, three, or four bursts of noise. When the experimenters substituted light flashes for the noise bursts they found no evidence of transfer. These results, combined with the results of the Fernandes and Church (1982) study suggest that abstract representations in rats may be confined to simple tasks requiring only a “less” and “more” judgment.

The monkeys in our lab have had extensive experience with many types of numerical tasks, but it is unclear whether the numerical knowledge gained from one task transfers to different types of tasks. The goal of this study was to investigate whether the monkeys in our laboratory possess a broad concept of number that includes different types of tasks and different presentation methods.

In Experiment 1, monkeys were trained to make one response after viewing a sequence of three circles flashed on a computer screen and another response after viewing a sequence of seven circles flashed on a computer screen. The monkeys were then presented with nonreinforced probe trials consisting of three or seven simultaneously visible circles. The goal was to assess whether or not the monkeys transferred the numerical knowledge gained in the sequential task to the simultaneous task by spontaneously providing a "three" response when presented with three simultaneously visible circles and a "seven" response when presented with seven simultaneously visible circles. Evidence of transfer would suggest that the monkeys possess an abstract representation of number that is not tied to a specific mode of presentation.

In Experiment 2, a different transfer paradigm was employed to investigate number concept in the same group of monkeys. In this experiment, the monkeys received reinforced presentations of simultaneously visible circles. For half of the monkeys, the correct response when presented with three simultaneously visible circles was the same as the correct response when presented with three sequentially presented circles in the prior experiment. Similarly, the correct response when presented with seven simultaneous circles was the same as the correct response when presented with seven sequential circles. For the other half of the monkeys, the reward contingencies were reversed so that the correct response for three simultaneous circles was the same as the correct response for seven sequential circles and the correct response for seven simultaneous circles was the same as the correct response for three sequential circles. If the monkeys transfer numerical knowledge from the sequential to the simultaneous task then it should be more difficult for the group with the reversed reward contingencies to learn this task than the group for which the reward contingencies stayed the same. Together, these two experiments will shed light on whether or not the monkeys in our laboratory have an abstract concept of number that spans different contexts and methodologies.

Experiment 1

Method

Participants. Eight male rhesus monkeys (*Macaca mulatta*) participated in this experiment. The monkeys, Gale (age 24 years), Hank (age 24 years), Willie (age 22 years), Murph (age 14 years), Chewie (age 6 years), Luke (age 6 years), Obi (age 2 years), and Han (age 3 years) were housed individually at the Language Research Center of Georgia State University according to federal animal housing standards and were not be food or water deprived during this study.

All of these monkeys are joystick-trained and have participated in previous computerized tasks related to various areas of cognitive research such as attention, metacognition, and concept learning (e.g., Flemming, Beran, & Washburn, 2007; Smith, Shields, & Washburn, 2003; Washburn & Gulledge, 2002). In addition, they have participated in a variety of tasks focusing on numerical ability. Some of these monkeys have previous experience with simultaneous and sequential numerical tasks, but none have experience with the specific sequential or simultaneous task used in the present experiment.

The older monkeys, Hank, Gale, Murph, and Willie have participated in a task requiring them to compare Arabic numerals to arrays of one through nine uniformly sized circles. Three of these monkeys were reinforced with a number of pellets proportional to the value of the stimulus chosen (i.e., four pellets for choosing the numeral 4 and three pellets for choosing an array of three circles). The fourth monkey, Hank, was reinforced with one pellet for choosing the stimulus with the highest value. All of the monkeys quickly learned to choose the stimulus with the highest numerical value (Harris, Gulledge, Beran, & Washburn, 2008). The current study also employed visual arrays of circles, but the circles varied widely in size. Also, the current task did not require the monkeys to compare two numerical stimuli. Instead, the monkeys learned to perform one response in the presence of three circles and another response in the presence of seven circles.

Three of the monkeys, Gale, Hank, and Willie have also participated in a task in which they learned to enumerate their sequential runs through a computerized maze and to choose the Arabic numeral or group of polygons that matched the number of runs in each series (Harris, Washburn, Beran, & Sevcik, 2007). The present study also involved a sequential task, but the monkeys were required to enumerate sequentially presented visual stimuli instead of their own sequential motor movements. Thus, specific knowledge about the sequential maze task was not applicable to the present study.

One of the monkeys, Murph, participated in a study that required him to watch as a computerized hand dropped items, one-at-a-time, into a cup on the computer screen (Beran, 2007). Although this task required Murph to enumerate sequential visual stimuli, he did not acquire any experience associating a sequence of items with an array of simultaneously presented items.

Apparatus. The monkeys were tested in their home cages using the LRC Computerized Test System (Rumbaugh, Richardson, Washburn, Savage-Rumbaugh, & Hopkins, 1989), which consists of a joystick attached to a computer and color monitor. The monkeys moved the joystick to control the movement of the cursor on the screen. The computer program recorded the stimulus displayed, the total size of the circles, the monkey's response, and the response time on each trial.

Stimuli. This experiment utilized both sequentially presented and simultaneously presented stimuli. The sequentially presented stimuli consisted of solid black circles flashed one at a time on a white background. The size of the circles ranged from 4 to 23 mm in diameter and the size of each circle was randomly chosen by the computer program before it was presented. Each time a circle flashed on the screen the computer program randomly assigned it to appear in one of 16 locations within the outline of a 100 mm × 100 mm square. Each trial consisted of three or seven sequentially presented circles. The numbers three and seven were chosen based on evidence from several studies that rhesus monkeys perform well above chance when comparing stimuli of a similar ratio (Beran, 2007; Harris et al., 2007; Harris et al., 2008).

The amount of time each circle was visible on the screen ranged from 200 to 700 ms and the inter-stimulus interval between the circles ranged from 100 to 350 ms. On each trial, the computer program randomly chose the presentation duration for each circle and the inter-stimulus intervals. This ensured that the total duration of the sequence and the rate of presentation varied from trial to trial. A sequence of three circles could range in total duration from 800 to 2,800 ms and a sequence of seven circles could range in total duration from 2,000 to 7,000 ms.

The simultaneously visible displays also consisted of three or seven circles. The circles used in the simultaneous displays were slightly smaller than those used in the sequential displays (2 - 21 mm rather than 4 - 23 mm in diameter) to avoid potential overlap on the screen, but they were otherwise identical. As was true for the sequential displays, the computer program randomly chose the size of the circles on every trial. At the start of each trial, the computer program randomly assigned each circle in the visual array to one of 16 locations within a square identical to the one used in the simultaneous displays. The circles all appeared on screen at the same time and remained visible until the monkey made a selection.

The third type of stimulus used in this experiment was an abstract red and blue shape approximately 75 mm × 75 mm. This shape was presented as a possible choice on every trial and its appearance remained constant throughout the experiment.

Training. The goal of training was to teach the monkeys to make one response after viewing three sequentially presented circles and another after viewing seven sequentially presented circles. Two of the four younger monkeys (Obi, Han, Chewie, and Luke) were randomly assigned to Group 1 and the other two were assigned to Group 2 because none of these monkeys had experience enumerating sequential items or events. Two of the four older monkeys were then randomly assigned to Group 1 and the other two were assigned to Group 2.

The target stimulus for Group 1 was three sequential circles and the target stimulus for Group 2 was seven sequential circles. This means that if the sequence consisted of three circles the monkeys in Group 1 were rewarded for moving the cursor into contact with the square outline in which the circles had appeared and the monkeys in Group 2 were rewarded for choosing the abstract shape. Conversely, if the sequence consisted of seven circles then the monkeys in Group 1 were rewarded for choosing the abstract shape and the monkeys in Group 2 were rewarded for choosing the square outline. For each trial, the computer program randomly determined whether three or seven circles would be presented. Therefore, approximately half of the sequences consisted of three circles and half consisted of seven circles.

In order to initiate a trial the monkeys were required to move the cursor into contact with a blue rectangle. After the rectangle was contacted it disappeared and the abstract shape appeared in the upper left corner of the screen and the square outline appeared in the upper right corner of the screen. Immediately after the square outline appeared, the sequential presentation of circles began. When the sequence was complete, a cursor appeared in the center of the bottom half of the screen and the monkey was allowed to choose one of the two stimuli.

Correct choices were rewarded with a melodic tone and the automatic delivery of one 94-mg fruit-flavored pellet. Incorrect responses resulted in a negative buzzing sound and a 10 second time-out, during which the computer screen remained blank. After the monkey received his reward or time-out, the blue rectangle again appeared in the center of the screen, allowing the monkey to initiate a new trial. Correction trials were utilized to prevent the monkeys from developing a side-bias in which they persisted in selecting the abstract shape or the square outline where the circles had appeared. This means that every incorrect trial was followed by a trial in which the same number of circles was presented. The number was repeated until the monkey made a correct response.

The monkeys were allowed continuous access to the task for several hours a day, several days a week. Trials were divided into 100 trial blocks for analysis and training for each monkey was complete when he reached an accuracy criterion of 80% correct for the three most recent blocks. Progress was assessed at the end of each day's training session and most of the monkeys reached criterion in the middle of a training session. Thus, most of the monkeys received several additional blocks of trials after reaching criterion and before the session ended.

Testing. The goal of testing was to determine whether the monkeys would transfer the responses learned during the sequential task to a simultaneous task. Novel simultaneous trials were randomly interspersed with the familiar sequential trials used during training. The simultaneous trials accounted for approximately 20% of the total number of trials. On the sequential trials, the computer program randomly determined whether three or seven circles would be presented so three circles were presented on approximately half of the simultaneous trials and seven circles were presented on the other half. The monkeys initiated the simultaneous trials in the same way as the sequential trials, by contacting the blue rectangle. After the rectangle was contacted it disappeared and the abstract shape appeared in the upper left corner of the screen, as was true for the sequential trials, and the square outline containing circles appeared in the upper right corner of the screen. The cursor appeared in the center of the bottom half of the screen at the same time as the other stimuli appeared.

The sequential trials continued to be scored and reinforced in the same manner they were during testing. The simultaneous trials were not reinforced. After a monkey made a selection on a simultaneous trial the circles and abstract shape immediately disappeared and were replaced by the blue rectangle, which allowed the monkey to initiate a new trial. The monkeys did not receive correction trials during the testing phase of the experiment.

Although the simultaneous trials were not reinforced in any way, responses were scored as correct or incorrect by the computer program. If the trial consisted of three simultaneous circles, Group 1 received a correct score for choosing the visible circles and Group 2 received a correct score for choosing the abstract shape. Conversely, if the trial consisted of seven simultaneous circles, Group 1 received a correct score for choosing the abstract shape and Group 2 received a correct score for choosing the visible circles. This means that the target number for Group 1 was always three, regardless of whether the trial was sequential or simultaneous and the target number for Group 2 was always seven, regardless of the mode of presentation.

The monkeys were allowed continuous access to the testing task for several hours a day, several days a week. Testing ended after the monkeys completed 2,000 trials, which equaled approximately 1,600 sequential trials and 400 simultaneous trials.

Analysis. As stated previously, when a monkey gave an incorrect response during training the trial was repeated until a correct response was given. The first response was included in the analyses, and all correction trials were excluded. Correction trials were not used during testing because the monkeys had already achieved accuracy criterion for the sequential trials.

The monkeys were not restrained in any way during this task and occasional disengagement in the middle of a trial resulted in unrealistically long trial times. All trials with response times longer than 10 seconds were excluded from analysis because this was approximately three standard deviations above the average trial time. This resulted in the exclusion of an average of 1.93% of training trials and an average of .44 % of testing trials across all eight monkeys.

Results

Training. The monkeys required between 8 and 67 blocks of training trials before reaching the accuracy criterion of 80% correct for the three most recent blocks. The number of blocks required for each monkey to reach criterion is shown in Table 6.1. A t-test comparing the accuracy of Group 1 and Group 2 revealed no significant difference, $t(6) = 1.50$, $p = .18$. This means that the target stimulus (three or seven) did not have a significant effect on performance.

The size of the circles varied within a sequential trial so that a sequence of three or seven circles usually consisted of a range of small and large circles. Despite this variation, when the areas of all circles in the sequence were summed together, sequences of seven circles consisted of a slightly larger average area compared to sequences of three circles. Across all eight monkeys, the average total area was 61.77 cm for sequences of three circles and 84.28 cm for sequences of seven circles.

To investigate the possibility that the monkeys were using the total area of the circles in each sequence instead of the number of circles as a cue to performing this task, the trials were divided into fifteen categories based on the total area of the circles. All size categories contained trials within a 10 cm range, with the smallest category consisting of trials with a total area of 15 - 25 cm and the largest category consisting of trials with a total area of 156 - 165 cm. The average performance across all eight monkeys as a function of the size category for trials consisting of three or seven circles is presented in Figure 6.1. In general, accuracy for three circles remained relatively stable across size categories, with the exception of a decrease in accuracy for the largest size category. It must be noted, however, that the largest category for 3 circles and 7 circles and the smallest category for 7 circles contained less than 15 data points for each of the monkeys so accuracy data for those categories may not be fully representative of the monkeys' abilities. Accuracy for seven circles was also relatively stable, with a slight increase for the most extreme size categories. The monkeys did not show sharp decreases in accuracy when the total size of three circles exceeded the average size for seven circles or when the total size of seven circles dropped below the average size for three circles, as would be expected if the monkeys were relying on the size of the circles to perform this task. These results indicate that size of the circles was not the primary cue that the monkeys used to perform this task.

Testing. All eight monkeys performed above chance on the approximately 1,600 familiar sequential trials presented during testing ($p < .05$, binomial sign test). It should be noted, however, that accuracy for half of the monkeys dropped below the 80% criterion level required to end the training phase. Thus it appears that the addition of nonreinforced simultaneous trials into the task disrupted performance on the sequential trials to some extent. Accuracy levels for the familiar sequential trials were as follows: Hank = 66.73%, Murph = 87.79%, Chewie = 87.90%, Luke = 88.57%, Han = 72.90%, Obi = 78.03%, Gale = 80.63%, Willie = 79.56%.

Seven of the eight monkeys performed at chance levels on the approximately 400 nonreinforced simultaneous trials presented during testing ($p > .05$, binomial sign test). Accuracy levels for those seven monkeys were as follows: Hank = 53.211%, Murph = 44.81%, Chewie = 50.68%, Luke = 46.43%, Han = 56.56%, Obi = 53.78%, Gale = 49.06%. All of these monkeys showed a strong bias for choosing either the circles or the abstract shape on every simultaneous trial regardless of the number of circles presented, which no doubt contributed to their low levels of accuracy. Murph, Chewie, Luke, and Han chose the circles on 91.39%, 81.64%, 90.93%, and 78.13% of simultaneous trials respectively. In contrast, Hank, Obi, and Gale exhibited a bias for the abstract shape and chose it on 70.05%, 77.76%, and 98.93% of simultaneous trials respectively.

The eighth monkey, Willie, performed above chance levels by correctly completing 59.84% of the nonreinforced simultaneous trials ($p < .01$, binomial sign test). In other words, he tended to make the same response to three circles regardless of whether they were simultaneously visible or sequentially presented and the same response to seven circles regardless of the presentation mode. Unlike the other seven monkeys, he did not show a strong bias for the circles or the abstract symbol, choosing the symbol on 56.60% of simultaneous trials. Thus, it appears that when Willie was confronted with novel simultaneous quantities, he was spontaneously able to transfer the numerical rules he had learned for sequentially presented circles to the novel and nonreinforced simultaneous task.

In a previous experiment involving nonreinforced probe trials, monkeys from our laboratory were able to maintain a high level of performance when presented with 500 nonreinforced trials interspersed every fifth trial with familiar reinforced trials. It was only during a second round of 500 nonreinforced trials that accuracy for the nonreinforced trials declined, which we speculated was due to decreased motivation for performing trials that were not associated with a reward. Thus, it seemed likely that the monkeys would be able to maintain a high level of motivation while performing all 400 nonreinforced trials in this experiment. In order to investigate the possibility that accuracy for the nonreinforced probe trials was higher at the beginning of the testing phase, accuracy was determined for each monkey after 250 testing trials (approximately 50 probe trials), 500 testing trials (approximately 100 probe trials), and 1000 testing trials (approximately 200 probe trials). First-trial performance for the simultaneous trials was also investigated to determine accuracy before the monkeys learned that the simultaneous trials would not be reinforced.

Analysis of first-trial accuracy revealed that four out of the eight monkeys responded correctly to the first simultaneous trial. This 50% overall accuracy rate provides no indication that the monkeys, as a group, immediately generalized to the new task. The results of the accuracy analysis at four different points during testing are presented in Figure 6.2. Willie's performance was above 60% for the first 50 probe trials and it remained relatively high throughout testing. In fact, his performance at 500, 1000, and 2000 trials exceeded chance levels of responding ($p < .05$, binomial sign test). Hank, Murph, Luke, and Gale were consistently at chance throughout testing. Chewie's performance was never significantly above chance, but he performed at 57.14% ($p = .39$) accuracy for the first 50 trials before his performance declined to levels closer to 50%. Han's performance was above 60% for the first two accuracy measurements and his performance of 61.96% after 500 trials was significantly above chance ($p < .05$). Despite high accuracy at the beginning of testing, his performance was not above chance for the last two accuracy measurements. Obi exhibited the highest evidence of spontaneous transfer at the beginning of testing. He performed at 69.77% accuracy for the first 250 trials, which was significantly above chance ($p < .05$), but his performance fell to chance level by the end of testing.

Discussion

All eight monkeys reached the accuracy criterion during training, although the number of trials required by each monkey varied widely. The average total size for seven circles was slightly larger than the average total size for three circles, which means that it was possible for the monkeys to rely on total size of the circles to perform the task. Further analyses revealed, however, that the monkeys were still able to perform the task when the total size of three circles exceeded the average size of seven circles and the total size of seven circles dropped below the average size of three circles. Thus, it appears that size of the circles was not the primary decision-making cue for the monkeys.

During testing, four of the monkeys, Hank, Murph, Luke, and Gale performed at chance levels on the simultaneous task and therefore showed no evidence of spontaneous transfer from the previous task. Han, Obi, and Chewie performed at high levels during the beginning of testing, but their performance declined towards the end of testing. This suggests that they may have initially been using numerical knowledge learned during the sequential task to perform the simultaneous task, but lost motivation after receiving a large number of nonreinforced trials and resorted to a simpler strategy of always selecting the visible circles or the abstract shape on simultaneous trials. The eighth monkey, Willie, exhibited high levels of accuracy on probe trials throughout testing, which indicated that he was able to generalize his learned responses for three and seven sequentially presented circles to three and seven simultaneously presented circles.

Experiment 2

Method

Participants, Apparatus, and Stimuli. In Experiment 2, the same participants, apparatus, and stimuli were utilized as Experiment 1.

Training. The transfer paradigm used in this experiment required the monkeys to have equal amounts of training on the sequential task. The monkeys had all completed a different number of sequential trials during Experiment 1 so the first step was to give the monkeys the necessary number of training trials to ensure that they all had equal experience.

During training the monkeys received sequential trials identical to those used in the training phase of Experiment 1. The sequential trials were reinforced in the same manner as they were during Experiment 1. The monkeys in Group 1 continued to receive food reinforcement for choosing the square outline after viewing three sequential circles and the abstract shape after viewing seven sequential circles and Group 2 continued to receive food reinforcement for making the opposite choices on the sequential trials. Correction trials were used during this phase of training, as they had been in Experiment 1 training.

The number of trials each monkey received depended on the number of sequential trials he had completed during the first experiment. During this training phase, the monkeys received enough sequential trials to equal 100 blocks (10,000 trials) total across both experiments. This included the 16 blocks of sequential trials completed during the testing phase of Experiment 1 and any additional blocks of trials the monkey may have received after reaching the accuracy criterion in the training phase of Experiment 1. For example, in Experiment 1 Hank required 67 blocks of sequential trials to reach criterion, received an additional 6 blocks after reaching criterion, and completed 16 blocks during testing. Thus, he received 11 blocks of sequential trials in this training phase (100 – (67 + 6 + 16) = 11). The blocks of sequential trials received by the other monkeys in this phase of training were as follows: Murph = 12, Chewie = 52, Luke = 47, Han = 44, Obi = 71, Gale = 42, and Willie = 66.

Testing. During testing, the monkeys were presented with trials involving three or seven simultaneously visible circles. These trials were identical to the nonreinforced simultaneous trials used in the testing phase of Experiment 1, except reinforcement was provided during this phase of the experiment. For half of the monkeys (two of the monkeys from Group 1 and two of the monkeys from Group 2), the correct response after viewing three simultaneous circles was the same as the correct response after viewing three sequentially presented circles in Experiment 1. Similarly, the correct response for seven simultaneous circles was the same as the correct response for seven sequential circles. This group will be henceforth referred to as the "No Switch" group. For the other half of the monkeys, the reward contingencies were reversed so that the correct response after viewing three simultaneous circles was the same as the correct response after viewing seven sequential circles and the correct response after viewing seven simultaneous circles was the same as the correct response after viewing three sequential circles. For example, Murph was in Group 1 during Experiment 1 and the reversed reward contingency group in this testing phase. Thus, when he was presented with three sequential circles in Experiment 1 or 2 he was rewarded for making contact with the square outline where the circles had been located. However, when he was presented with three simultaneously visible circles in the testing phase of Experiment 2 he was rewarded for making contact with the abstract shape. Making contact with the square outline containing three visible circles was scored as incorrect and resulted in a buzzing sound and a time-out. The reversed reward contingency group will be henceforth referred to as the "Switch" group.

Correct and incorrect testing trials resulted in the same feedback as correct and incorrect training trials. Correction trials were used as they had been in training. The monkey was given the opportunity to initiate a new testing trial after receiving his reward or time-out for the previous testing trial.

The monkeys were allowed continuous access to the task for several hours a day, several days a week. The trials were administered in 100 trial blocks and testing was complete for each monkey when he had reached an accuracy criterion of 80% correct for the three most recent blocks.

Analysis. As was true in Experiment 1, all correction trials and all trials with response times longer than 10 seconds were excluded from analysis. The response time filter resulted in the exclusion of an average of 1.01% of training trials and an average of .13 % of testing trials across all eight monkeys.

Results

Training. All monkeys exhibited a high level of accuracy that exceeded, or came very close to exceeding, the accuracy criterion for training in Experiment 1. Hank performed at 79.64% accuracy, Murph at 88.42%, Chewie at 94.12%, Luke at 93.87%, Han at 86.64%, Obi at 90.70%, Gale at 86.55%, and Willie at 89.02%.

Testing. The goal of testing was to compare the performance of the monkeys in the No Switch group to the monkeys in the Switch group, which could be affected by the ability of the monkeys to learn a new numerical task and to perform the training task in particular. To rule out these confounding variables the number of blocks required to reach criterion during the training task in Experiment 1 and the accuracy of the monkeys when performing the training task in Experiment 2 were compared for the Switch and No Switch groups. The No Switch group required an average of 34 blocks to reach criterion in Experiment 1, which was not significantly different than the average 30.5 blocks required by the Switch Group, $t(6) = .21$, $p = .84$. In addition, the average accuracy for the No Switch group in the training phase of this experiment was 87.71%, which was not significantly different than the average accuracy of 89.53% for the Switch group, $t(6) = .52$, $p = .62$.

The number of trial blocks (100 trials per block) that each monkey required to reach the accuracy criterion is presented in Table 6.2. On average, the No Switch group required 12 blocks and the Switch group required 10.75 blocks of trials. The target number for the monkeys varied within the Switch and No Switch groups depending on the group to which they were originally assigned in Experiment 1. For instance, Murph was assigned to Group 1 during the first experiment so his target number during the training phase of both experiments was three. Murph was then assigned to the Switch group so his target for the testing phase of this experiment was seven. Gale was also assigned to the Switch group, but he was originally in Group 2 so his target for the testing phase of this experiment was three. A two-way ANOVA of the effects of original group and new group on the number of trials required to reach criterion was performed with data from all eight monkeys. *Original group* refers to whether the monkeys were in Group 1 or Group 2 during training and *new group* refers to whether the monkeys were in the No Switch or Switch group during testing. This analysis yielded no significant differences (original group, $F(1, 4) = 1.67$, $p = .27$; new group, $F(1, 4) = .07$, $p = .81$; original group × new group interaction, $F(1, 4) = 4.93$, $p = .09$).

A rank ordering of the individual monkeys according to the number of trials required to reach criterion also did not reveal strong evidence of a group effect. The order of the monkeys from least number of trial blocks to most number of trial blocks required is as follows: Obi (No Switch), Murph (Switch), Han (No Switch), Luke (No Switch), Willie (Switch), Chewie (Switch), Gale (Switch), and Hank (No Switch). Luke, a monkey from the No Switch group, and Willie, a monkey from the Switch group, were tied for the middle position at 12 blocks of trials. Two monkeys from the No Switch group and one monkey from the Switch group required fewer trials than Luke and Willie and two monkeys from the Switch group and one monkey from the No Switch group required more trials than Luke and Willie.

Although the No Switch group did not reach criterion in fewer trials, as would be expected if the monkeys were transferring numerical knowledge from the sequential task to this simultaneous task, spontaneous transfer may have occurred at the very beginning of testing before the monkeys had a chance to learn from repeated reinforcement. Results from the first five and ten probe trials for each monkey are presented in Table 6.3. There were no significant differences between the accuracy of the No Switch and Switch groups for these initial probe trials, first five trials: $t(6) = 1.46$, $p = .19$ and first ten trials: $t(6) = 1.12$, $p = .31$. However, the No Switch group did have a higher mean accuracy for the initial probe trials than the Switch group. The No Switch group performed at 80% and 62.5% accuracy for the first five and first ten probe trials respectively, while the Switch group performed at 55% and 47.5% for those same probe trials. Only the 80% accuracy level was significantly greater than chance ($p < .05$).

Discussion

During training, all monkeys exhibited a high level of accuracy that exceeded, or came very close to exceeding, the accuracy criterion for training in the first experiment. The Switch and No Switch groups did not differ significantly in the average number of blocks required to reach criterion during the training task in Experiment 1 or the average accuracy for the training task in Experiment 2. Thus, the two groups exhibited an equal ability to learn a new numerical task and to perform the training task in particular.

If the monkeys transferred numerical knowledge from the sequential task to the simultaneous testing task, the No Switch group would be expected to reach the accuracy criterion in fewer trials than the Switch group. Results revealed that the No Switch and Switch groups did not differ significantly in the average number of trial blocks required to reach the accuracy criterion. Spontaneous transfer that may have occurred at the very beginning of testing before the monkeys had a chance to learn from repeated reinforcement was also investigated by looking at the initial probe trials. The average accuracy for the No Switch group was higher than that of the Switch group on the first five and first ten probe trials, but not significantly so.

General Discussion

All eight monkeys learned to make one response after viewing a sequence of three dots flashed on a computer screen and another response after viewing a sequence of seven dots. The total size of the circles presented in each trial was somewhat correlated with number, so it was possible for the monkeys to use the size of the circles to perform some of the trials correctly. A post-hoc analysis revealed that extreme size did affect performance, but that size was not the primary decision-making cue for the monkeys.

The average total duration for a sequence of three circles was longer than the average total duration for seven circles, but duration and rate of presentation for each circle varied from trial to trial. This means that it was possible for the monkeys to use temporal variables to help them perform the task, but these temporal variables were complicated and unreliable cues. Even if the monkeys were using temporal variables to perform the sequential task, temporal knowledge could not help them perform the simultaneous tasks. Given that the monkeys were not relying on size cues to perform the sequential task and could not use temporal variables to aid them in the simultaneous task, any evidence of transfer between the two tasks was most likely due to transfer of numerical knowledge.

Results from the testing phase of Experiment 1 revealed that four out of the eight monkeys showed no evidence of transfer from the sequential training task to nonreinforced simultaneous probe trials. Instead of spontaneously providing a “three” response when presented with three simultaneously visible circles and a “seven” response when presented with seven simultaneously visible circles, they each developed a strong bias for contacting the numerical stimulus or the abstract shape, regardless of the number of circles presented. The sequential and simultaneous trials were made as similar as possible to facilitate transfer, but despite this effort it is possible that the monkeys viewed them as two completely separate tasks. If that is the case, it is not surprising that they did not use the same numerical rules to perform the sequential and simultaneous trials.

In contrast, three of the monkeys performed at high levels on the simultaneous probe trials during the beginning of testing, but were unable to sustain performance across all 400 probe trials (2,000 trials in total). This suggests that they may have initially been using the reward contingencies learned during the sequential task to perform the simultaneous task, but lost motivation after receiving a large number of nonreinforced trials. The eighth monkey, Willie, exhibited high levels of accuracy on probe trials throughout testing, which indicated that he was able to generalize his learned responses for sequentially presented circles to the same number of simultaneously presented circles. Although this task proved difficult for most of the animals, Willie's success provides evidence that at least some rhesus monkeys have a capacity for numerical transfer.

Only two quantities, three and seven, were used in this task so it could be performed by representing the number of sequential and simultaneous circles simply as "few" and "many" without representing the number of circles as a specific quantity. This means that the monkeys could perform the simultaneous task without utilizing precise quantity information from the sequential task. Despite this possibility, the only way to discriminate between a large and small quantity when perceptual and temporal cues are discounted is numerical information. The monkeys must recognize that "few" and "many" differ in number in order to discriminate between these two choices. Thus, some of the monkeys were capable of numerical transfer, regardless of whether the monkeys were representing the numbers precisely.

In the second experiment, the monkeys were presented with reinforced simultaneous trials and the number of trials required for the monkeys to learn the task was measured. The goal was to compare the performance of the No Switch group for which numerical contingencies remained the same regardless of presentation method (sequential or simultaneous), with performance of the Switch group in which numerical contingencies were reversed from the sequential task to the simultaneous task. If the monkeys were categorizing sequentially and simultaneously presented stimuli together on the basis of number, the Switch group should have required more trials to learn the simultaneous testing task. Results revealed, however, that the two groups required approximately the same number of trials to reach the accuracy criterion.

The fact that Willie was performing above chance on simultaneous trials at the end of Experiment 1 likely affected his performance when he was presented with identical simultaneous trials in Experiment 2. Willie was in the Switch group, however, which means that previous proficiency with the task would most likely be a detriment to his performance in Experiment 2. This did not appear to be the case, as Willie required the same number of trials to learn the task as the No Switch group required on average. This finding is interesting because it suggests that Willie did not transfer the numerical rule used during the testing phase of Experiment 1 to the testing phase of Experiment 2, even though the simultaneous trials were identical.

Although the No Switch group did not reach criterion in fewer trials, as would be expected if the monkeys were transferring numerical knowledge from the sequential task to this simultaneous task, spontaneous transfer may have occurred at the very beginning of testing before the monkeys had a chance to learn from repeated reinforcement. Analysis revealed that the average accuracy for the No Switch group was 80% for the first five probe trials and 62.5% for the first ten probe trials, which was higher than that of the Switch group, although not significantly so. Thus, neither the overall accuracy nor the accuracy on initial probe trials provided strong evidence that the numerical rules learned during the sequential task were affecting performance on the simultaneous task.

Overall, these experiments indicate that some rhesus monkeys have an abstract concept of number that reaches across presentation mode. These findings compliment the previous findings by Jordan et al. (2005) that rhesus monkeys possess an abstract concept of number that extends across two different sensory modalities. The results of this study also indicate, however, that an abstract concept was not automatically activated in all numerical situations. In the first experiment, there was some evidence that monkeys viewed three sequential circles as similar to three simultaneous circles and seven sequential circles as similar to seven simultaneous circles and responded based on previously learned reward contingencies for those numbers. These same monkeys failed to show the same capacity in the second experiment.

The fact that the monkeys typically work on several different tasks each day might have interfered with any natural tendency to transfer knowledge from one task to another, regardless of how similar the tasks appeared. In that were the case, the monkeys may have recognized that three sequentially presented circles and three simultaneously presented circles were in the same numerical category, without responding to both in the same manner. Thus, the study may have failed to capture the full potential of the monkeys in terms of numerical transfer. It is also possible that monkeys, unlike humans, cannot easily abstract number across different contexts, and therefore the ability is only exhibited by some monkeys under limited conditions.

These remaining questions regarding numerical abstraction also leave open the question of whether or not rhesus monkeys behave in a way that fulfills all five of the counting principles proposed by Gelman and Gallistel (1978) and the definition of true counting proposed by Davis and Pérusse (1988). It seems likely that numerical abstraction, like numerical ability in general, is not “all or none”, but rather a graded capacity that exists in nonhuman primates to a lesser extent than in adult humans with a lifetime of mathematical training. If that is the case, devising an artificial threshold over which animals must pass before they are declared to possess numerical abstraction is less important than investigating the extent of this ability in different animal populations and the conditions under which the capacity is demonstrated.

References

Beran, M. J. (2007). Rhesus monkeys (*Macaca mulatta*) enumerate large and small sequentially presented sets of items using analog numerical representations. *Journal of Experimental Psychology: Animal Behavior Processes, 33,* 42-54.

Church, R. M., & Meck, W. H. (1984). The numerical attribute of stimuli. In H. L. Roitblatt, T. G. Bever, & H. S. Terrace (Eds.), *Animal Cognition* (pp. 445-464). Hillsdale, NJ: Lawrence Erlbaum Associates.

Davis, H., & Albert, M. (1986). Numerical discrimination by rats using sequential auditory stimuli. *Animal Learning and Behavior, 14,* 57-59.

Davis, H., & Pérusse, R. (1988). Numerical competence in animals: Definitional issues, current evidence, and a new research agenda. *Behavioral and Brain Sciences, 11,* 561-615.

Fernandes, D. M., & Church, R. M. (1982). Discrimination of the number of sequential events by rats. *Animal Learning and Behavior, 10,* 171-176.

Flemming, T. M., Beran, M. J., & Washburn, D. A. (2007). Disconnect in concept learning by rhesus monkeys (*Macaca mulatta*): Judgment of relations and relations-between-relations. *Journal of Experimental Psychology: Animal Behavior Processes, 33,* 55-63.

Gelman, R., & Gallistel, C. R. (1978). *The child's understanding of number*. Oxford, England: Harvard University Press.

Harris, E. H., Gulledge, J. P., Beran, M. J., & Washburn, D. A. (2008). *What do Arabic numerals mean to macaques?* Manuscript submitted for publication.

Harris, E. H., Washburn, D. A., Beran, M. J., & Sevcik, R. A. (2007). Rhesus monkeys (*Macaca mulatta*) select Arabic numerals or visual quantities corresponding to a number of sequentially completed maze trials. *Learning and Behavior, 35*, 53-59.

Jordan, K. E., Brannon, E. M., & Gallistel, C. R. (2006). The multisensory representation of number in infancy. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 103,* 3486-3489.

Jordan, K. E., Brannon, E. M., Logothetis, N. K., & Ghazanfar, A. A. (2005). Monkeys match the number of voices they hear to the number of faces they see. *Current Biology, 15,* 1034-1038.

Rumbaugh, D. M., Richardson, W. K., Washburn, D. A., Savage-Rumbaugh, E. S., & Hopkins, W. D. (1989). Rhesus monkeys (*Macaca mulatta*), video tasks, and implications for stimulus-response spatial contiguity. *Journal of Comparative Psychology, 103,* 32-38.

Smith, J. D., Shields, W. E., & Washburn, D. A. (2003). The comparative psychology of uncertainty monitoring and metacognition. *Behavioral and Brain Sciences, 26,* 317-373.

Starkey, P., Spelke, E. S., & Gelman, R. (1983). Detection of intermodal numerical correspondences by human infants. *Science, 222,* 179-181.

Starkey, P., Spelke, E. S., & Gelman, R. (1990). Numerical abstraction by human infants. *Cognition, 36,* 97-127.

Washburn, D. A., & Gulledge, J. P. (2002). A species difference in visuospatial working memory in human adults and rhesus monkeys. *International Journal of Comparative Psychology, 15,* 288-302.

Table 6.1
The Number of Blocks (100 Trials Each) Required for Each Monkey to Reach the Accuracy Criterion in the Training Phase of Experiment 1

Group 1 (Target 3)		Group 2 (Target 7)	
Monkey	Blocks	Monkey	Blocks
Hank	67	Han	36
Murph	63	Obi	13
Chewie	22	Gale	29
Luke	20	Willie	8
Average	**43**	**Average**	**21.5**

Note. The difference between the average of Group 1 and Group 2 was not significant, $p > .05$

Table 6.2
The Number of Blocks (100 Trials Each) Required for Each Monkey to Reach the Accuracy Criterion in the Testing Phase of Experiment 2

No Switch		Switch	
Monkey	Blocks	Monkey	Blocks
Hank	29	Murph	4
Luke	12	Chewie	13
Han	4	Gale	14
Obi	3	Willie	12
Average	**12**	**Average**	**10.75**

Note. The difference between the average of the No Switch and Switch group was not significant, $p > .05$

Table 6.3
Percentage Accuracy for the First 5 and 10 Probe Trials in the Testing Phase of Experiment 2

No Switch			Switch		
Monkey	5 probe trials	10 probe trials	Monkey	5 probe trials	10 probe trials
Hank	100	70	Murph	80	70
Luke	80	60	Chewie	40	50
Han	80	80	Gale	80	50
Obi	60	40	Willie	20	20
Average	**80**	**62.5**	**Average**	**55**	**47.5**

Note. The difference between the No Switch and Switch groups was not significant for the first five probe trials or the first ten probe trials, $p > .05$

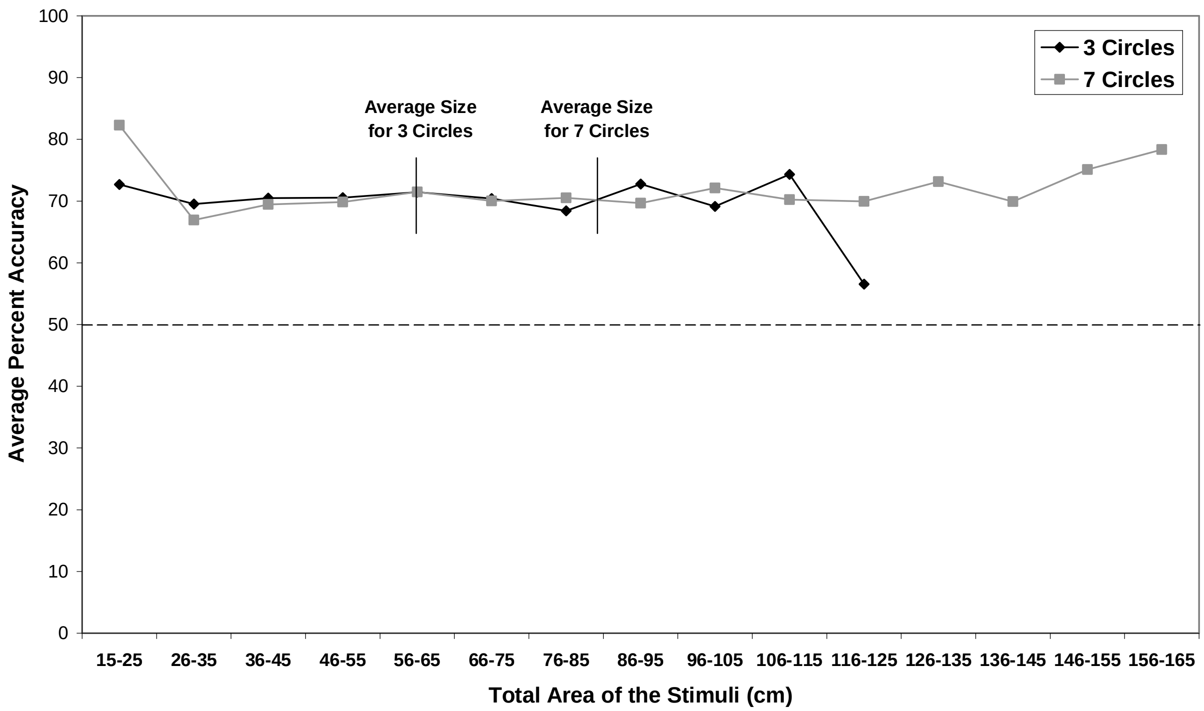


Figure 6.1. The mean performance across all eight monkeys in the training phase of Experiment 1 as a function of the total area of the stimuli presented.

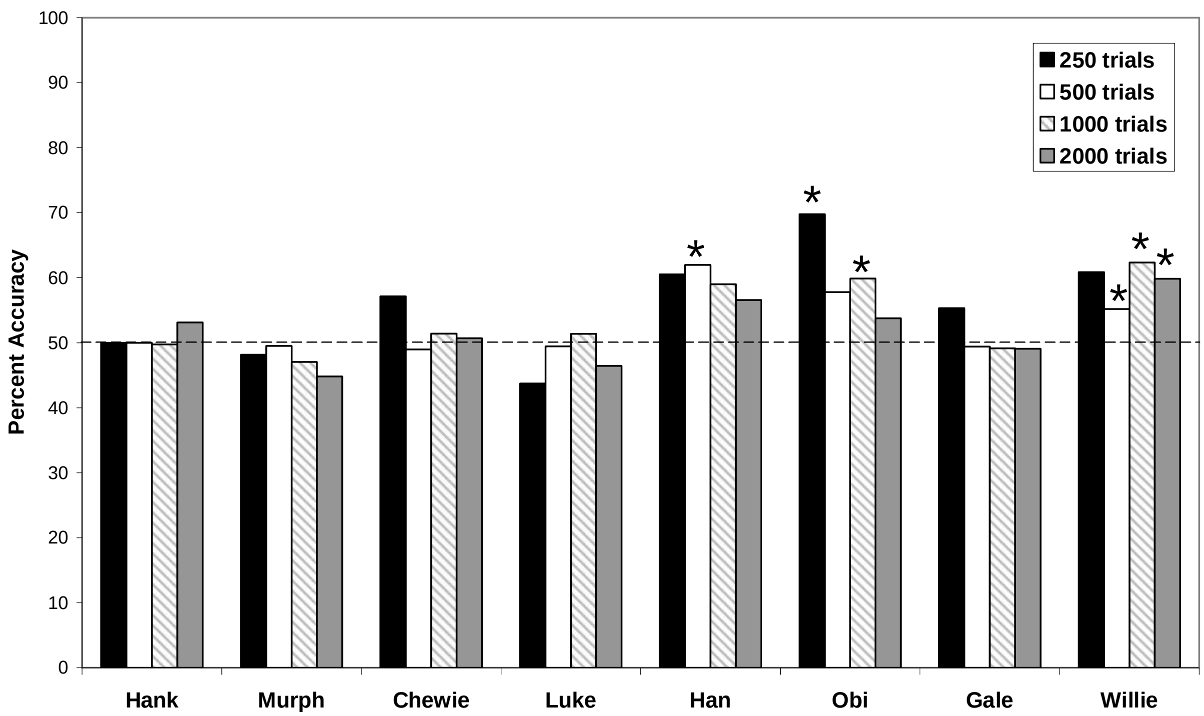


Figure 6.2. Accuracy for the nonreinforced simultaneous trials after 250, 500, 1000, and 2000 testing trials in Experiment 1. Horizontal dashed line denotes chance level of performance. Asterisks denote accuracy levels that are significantly above chance performance ($p < .05$).

Chapter 7: Discussion

Two decades ago, Davis and Pérusse (1988) provided an influential review of the comparative literature on numerical competence. Although their review served primarily to define terms and to identify methodological issues, it (and the commentaries that followed it) also articulated the theoretical debate that provides the foundation for the present study. One position is that animals' responsiveness to stimulus numerousness is best explained by psychologically fundamental mechanisms that reflect the animals' experience with those specific stimuli, including the rich perceptual array of nonnumerical cues correlated with number, and the consequences of responses to those stimuli. Another position is that an animal's numerical competence is a reflection of mental representations of categorical knowledge about number (albeit imprecise) that emerge from experience with quantities in various contexts, and that can guide behavior under novel task demands.

The studies included in this dissertation provided support for the latter perspective that monkeys possess a concept of number that is not based on simple associative or perceptual mechanisms. The primary source of evidence for this conclusion is performance in a series of control tests in which reliance on nonnumerical cues was not possible and transfer tests in which the animals were presented with novel combinations of numerical stimuli, novel reward contingencies, and novel response demands. In addition, the monkeys exhibited an ability to map arbitrary symbols, in the form of Arabic numerals, onto representations of analog quantities and to use the symbols in new and emergent ways. This provides evidence that Arabic numerals are not simply sign-stimuli associated with specific response-reward histories, but rather serve a symbolic function.

Based on these abilities it is tempting to ascribe to these monkeys a numerical concept identical to our own. However, when we look beyond the question of whether or not they can perform these numerical tasks and examine the underlying mental processes, species differences become clear. For example, Chapters 2 through 5 provided ample evidence that monkeys represent number in a less precise manner than humans do, as predicated by the analog magnitude model (Dehaene, 1997; Gallistel & Gelman, 1992). In addition, the mixed findings regarding transfer from a sequential to a simultaneous task in Chapter 6 cast doubt on whether these monkeys have the same abstract concept of number as adult humans.

The first step to determining whether monkeys have a concept of number and a symbolic understanding of Arabic numerals was to rule out nonnumerical variables that may be responsible for responding in numerical tasks. For example, monkeys may compare numerals using representations of cardinal values, or they may make the same comparisons using a learned matrix of 2-choice discriminations (e.g., pick the numeral 7 when presented with 6, but not when presented with 8) or hedonic values acquired during training (e.g., pick the numeral 7 when presented with 6 because it is more satiating).

The results from Chapter 2 revealed that monkeys were able to choose the largest value when presented with novel probe trials involving one numeral and one dot array. This was true even on the first exposure to these trials. This finding allowed us to rule out a matrix of learned values as the basis of responding. Additionally, results from one monkey receiving probabilistic reinforcement suggested that the monkeys were not solving the comparisons based on the hedonic value of the numerals. His responses were similar to the responses of the other monkeys, even when numerical value opposed hedonic value. For example, he was able to correctly respond to a comparison of four dots and the numeral 3, despite the fact that the numeral 3 had a much richer reinforcement history.

The fact that the monkeys were not responding to numerical pairs based on a complex matrix of memorized responses or hedonic values suggests that these monkeys understood that Arabic numerals represented absolute values that could be ordered and compared on a relative basis. It must be noted, however, that the monkeys appeared to be using approximate quantity information rather than exact. Performance suffered when the numerical distance between the numeral and dot quantity was small.

The evidence that monkeys are able to make comparisons based on approximate cardinal information is particularly interesting given that cardinal numerical information does not appear to be an ecological necessity for animals in the wild. It is important for a monkey to know which fruit tree in the forest contains the most fruit, but it is less important to know the exact number of fruit on any of the trees or the absolute amount by which the trees differ. It is also difficult to think of an example in which a monkey would need to associate a quantitative value with an abstract symbol.

This discrepancy between behaviors demonstrated in the lab and behaviors routinely observed in the wild brings up the issue of untrained behaviors versus trained behaviors. Monkeys in the wild are obviously not familiar with computer joysticks and procedures such as match-to-sample. However, they almost certainly have the same potential to learn these things as the monkeys in our laboratory. Although laboratory research captures behaviors that are beyond the scope of the normal survival behaviors studied using experimentally naïve monkeys in their natural habitats, both perspectives are necessary to understand numerical cognition in animals. Research conducted with experimentally naïve animals (e.g., Flombaum, Junge, &

Hauser, 2005; Hauser, Carey, & Hauser, 2000; Menzel, 1960; Santos, Barnes, & Mahajan, 2005; Uller, Hauser, & Carey, 2001), provides us with valuable information about cognitive adaptations, including those related to numerical processing, that produce survival benefits for those animals. In other words, we learn how the animals use cognitive processes, such as quantitative comparisons, to maximize food intake, avoid dangerous conflicts, and excel at other activities that are necessary for their continued existence. This research also provides us with a glimpse into our own possible hominid past. We can imagine how our modern system of number processing arose from environmental adaptations that improved our own evolutionary fitness.

In contrast to studies utilizing experimentally naïve monkeys, laboratory studies often involve experimentally savvy monkeys with years of training on numerical tasks. The main goal of these studies is not to document the everyday behaviors that the animal might exhibit in its natural habitat, but rather to challenge the animal in novel and controlled ways that will uncover the animal's natural learning potential. In other words, the focus is on the capacity to learn numerical tasks, rather than the numerical behaviors that are normally observed in untrained animals. By training monkeys on numerical tasks, just as humans are trained for years in math class, we can discover whether or not monkeys, like humans, are able to build on their innate mathematical foundations and succeed at more sophisticated tasks involving symbolic and absolute number knowledge. Both the successes and failures of the monkeys on these complex tasks provide us with information about the nature and limits of their mental numerical models.

Despite the benefits of laboratory study, the fact that many of the abilities observed in trained animals provide little survival advantage means that researchers must be especially careful to rule out other explanations before attributing performance to absolute numerical knowledge. For example, the ability of the monkeys to choose the largest value for probe trials involving one numeral and one dot array in Chapter 2 strongly suggested that the Arabic numerals represented absolute quantities, but there is an alternative possibility that should be mentioned. It is possible that performance was not based on the comparison of numerical representations, but instead reflected integration of two learned sequences. Previous research has shown that monkeys can learn to order two lists of arbitrary stimuli and immediately respond correctly at a greater than chance level when presented with comparisons of two items from different lists (D'Amato & Colombo, 1988; Terrace, Son, & Brannon, 2003). It is possible that the monkeys in this study perceived the numerals as one arbitrary list of stimuli and the dot quantities as another arbitrary list and correctly responded to pairs of numerals and dots using only knowledge of their ordinal position. In other words, monkeys may correctly order the same numerals using knowledge that the numeral 5 is followed by 4 is followed by three, or that 5 > 4 > 3, without regard for the cardinal values of the numerals or the absolute difference between them.

Thus, the second step in determining the nature of number concept in rhesus monkeys was to investigate the precision of the numerical representation by testing for ordinal versus absolute (cardinal) numerical knowledge. The ordinal tasks in Chapter 3 revealed that both monkeys learned to produce three 5-item lists at greater-than-chance levels. The monkeys showed no advantage when learning the list of Arabic numerals compared to the novel list of signs and the novel list of colors. In the second experiment the monkeys performed significantly above chance levels for all types of pair-wise comparisons with items from different lists, with no clear facilitative effects for comparisons involving numerals. These results indicate that the monkeys were responding on the basis of the newly learned ordinal positions of the items.

The fact that the monkeys showed no advantage when learning a list of Arabic numerals or integrating the numeral list with the lists of arbitrary signs and colors suggests that ordinal knowledge did not inhere in the representations formed during the presentation of Arabic numerals in the first experiment of Chapter 2. Thus, the idea that the monkeys in Chapter 2 were able to compare probe trials of dots and numerals by integrating two learned ordinal lists was not supported by these findings. However, it is also possible that the monkeys acquired ordinal knowledge during the experiment in Chapter 2, but that learning was lost between studies due to time and interference from intervening experimental tasks.

In addition to successfully comparing numerals, colors, and signs that were trained as ordered lists, one monkey was able to make ordinal comparisons using analog quantities. The monkeys had never received serial training involving lists of quantities, but when we presented analog quantities within the context of making ordinal judgments, one monkey spontaneously used the magnitude of the polygon set to determine its ordinal position relative to the learned-list stimuli.

In summary, the experiments in Chapters 2 and 3 strongly suggest that the monkeys possess a concept of number that includes both cardinal and ordinal value. In other words, the monkeys are not representing visual quantities as separate and unrelated categories in the same way that we would represent categories such as “tree” and “flower.” Instead, the monkeys understand that numerical categories have consistent ordinal relationships that allow them to be ordered and compared. Likewise, the monkeys are not representing visual quantities as items in an arbitrarily ordered list in the same way that humans would represent letters of the alphabet. Instead, they understand that visual arrays are ordered based on absolute numerical values. This knowledge allows the monkeys spontaneously to integrate visual quantities with arbitrary symbols that were learned as an ordered list. In addition, the monkeys learned through training that Arabic numerals with no inherent numerical meaning have symbolic value representing numerical quantities with both cardinal and ordinal properties. Thus, these monkeys could compare and order Arabic numerals in the same way as visual arrays.

Although these conclusions provide a fairly detailed picture of how the monkeys in our laboratory represent simultaneously presented Arabic numerals and visual arrays, the conclusions may not generalize to tasks involving sequential items or events. Experiments in Chapters 2, 4, and 5 were designed to assess the use of a number concept by rhesus monkeys in a series of sequential tasks.

In Chapter 4, we trained four Arabic numeral-experienced monkeys on a series of reinforced (R) and nonreinforced (N) computerized maze trials. During training on an RRRN series, two of the four monkeys developed a “slow, fast, faster, slow” pattern, which suggested they were anticipating the final nonreinforced trial. The monkeys initiated the trials themselves and no strict temporal controls were employed, but an analysis of the data made it clear that they were not using duration as a primary cue to predict when the nonreinforced trial would occur. The other two monkeys performed gradually slower on each trial in a series, which made it impossible to ascertain whether or not they were predicting precisely when the final trial would occur.

During testing, the monkeys receiving probe trials of the numerals 2 and 4 developed a pattern of performing more slowly on the nonreinforced trials than the reinforced trials, indicating the use of the changing target numeral cues to anticipate those final nonreinforced trials. The monkeys receiving probe trials of the numerals 2 through 8 did not use the changing numeral to predict precisely when the nonreinforced trial would occur in each series, but did perform faster overall on series with higher target numerals.

In the second experiment of Chapter 2, we presented monkeys with two Arabic numeral cues in a computerized maze and each numeral was “baited” with the corresponding number of pellets. We reasoned that a monkey could travel to the larger numeral the corresponding number of times and then behaviorally indicate that he knew he had exhausted the pellets at that location by traveling to the smaller numeral. In contrast, if the monkeys know only the ordinal and not absolute values corresponding to the numerals, then they would have no basis for knowing when to stop responding to the larger of the two numerals. This design allowed us to assess the monkeys’ understanding of the absolute value of the numerals in a sequential task.

For three of the four monkeys, responding was above chance levels and observed behavior was better simulated by an algorithm that selected numerals in proportion to their relative magnitudes than one that selected numerals by chance. In other words, the monkeys were twice as likely to select the numeral 4 rather than the numeral 2 when they were paired, versus having a .50 probability of selecting each numeral. However, even this simulation failed to capture the level of errorless trials that were observed with target proportions of 0.6 and greater. These data suggest that the monkeys had some understanding of the quantity symbolized by the numerals.

Despite performing at greater than chance levels, the monkeys tended to touch a target more times than was represented by the numeral. These errors suggested that the monkeys were not representing that quantity precisely. In other words, the monkeys were not enumerating exactly three responses to the numeral 3, but were instead responding in a more approximate manner. Overall, this study suggests that Arabic numerals provide more information to the monkeys than ordinality (which numeral is bigger) or even the relative magnitudes of the numerals, but do not provide them with exact quantity information.

In Chapter 5, we further investigated the precision of the monkeys' numerical representations by testing for absolute (cardinal) numerical knowledge in a sequential enumeration task. During the course of this study, all of the monkeys learned to match randomly intermixed series of one or nine maze trials with the correct Arabic numeral or visual quantity when tested with a same/different discrimination. Two of the monkeys achieved accuracies greater than 70% for the numerosities 1, 5, and 9, within the first 500 presentations, but they did not reach criterion when the numerosity 3 was added to the experimental set. Their ability to perform the task with three numerosities indicates that their representation of the maze trials went beyond a simple representation of "one" and "many." The fourth monkey in the study performed the task with randomly intermixed series of the numerosities 1, 3, 5, and 9, but failed to achieve the accuracy criterion after the numerosity 2 was added to the set.

The monkeys' error patterns were not related to the amount of time they spent on the maze trials in each series. The monkeys did not tend to choose numerosities that were higher than the correct choice after spending more time than usual on a particular series. This finding indicates that responding was based on the number of maze trials and not the duration of the maze trials, which provides evidence for a concept of number rather than a reliance on timing processes.

Interestingly, the monkeys performed equally well regardless of whether the numerical stimulus was an Arabic numeral or visual dot quantity. Although the visual quantities provided more inherent numerical information than the numerals, the monkeys have had a variety of testing experiences involving Arabic numerals. Their ability to match a series of maze trials to either a visual quantity or an Arabic numeral indicates flexibility in their performance strategy.

Together, the studies in Chapters 2, 4, and 5 suggest that the monkeys' numerical competence in sequential tasks, as in the simultaneous tasks, reflects a mental representation of approximate cardinal number. Although these animals, like adult humans, have the ability to keep track of their own sequential responses, alter their motor responses based on Arabic numeral cues, and match a number of sequential responses with the corresponding numeral or visual quantity, there is scant evidence that the monkeys generalized numerical learning in simultaneous tasks to these sequential tasks. Nonetheless, it is clear that they can perform both types of tasks with training so their numerical abilities are not limited to a specific context.

All of these studies demonstrated that the monkeys in our laboratory possess a concept of number that is not based on simple associative or perceptual mechanisms or experience with specific stimuli. This numerical concept guided behavior under a variety of novel task demands. In addition, these studies provided evidence that the monkeys understood that numerals, which are abstract symbols with no physical properties that correlate with the quantities they represent, nonetheless do represent specific numerical quantities.

When the data are inspected closely, however, it becomes clear that monkeys represent number in a less precise manner than humans, consistent with the analog magnitude model (Dehaene, 1997; Gallistel & Gelman, 2000; Meck & Church, 1983). Unlike adult humans, the monkeys were not able to take advantage of the numerical precision made possible by the use of Arabic numerals. Instead, the monkeys were able to perform the tasks at levels greater than chance by representing approximate, rather than exact, quantities.

For example, in the first experiment in Chapter 2 the monkeys were able to compare numerals and dot quantities, but performance suffered when the numerical distance between the numeral and dot quantity was small. This distance effect, which is commonly seen in animal numerical studies (e.g., Anderson, Stoinski, Bloomsmith, Marr, Smith, & Maple, 2005; Boysen & Berntson, 1995; Brannon & Terrace, 1998, 2000; Judge, Evans, & Vyas, 2005; Smith, Piel, & Candland, 2003), provides evidence that the monkeys were using continuous representations of magnitude rather than representations of exact number. Two magnitude representations, like two physical magnitudes such as length or weight, are more difficult to compare when they are close in distance than when they are far apart (Gallistel & Gelman, 1992; Nieder & Miller, 2004; Whalen, Gallistel, & Gelman, 1999).

A distance effect was also found in the ordinal learning experiment in Chapter 3 in which the monkeys compared numerals, colors, and abstract symbols from different lists. The monkeys performed better as the ordinal distance between the two comparison items increased. This remained true when the monkeys were given probe trials involving visual dot quantities that were not trained as lists. These results indicate that although the monkeys were responding on the basis of the ordinal position of the items, the ordinal knowledge was inexact (i.e., the monkeys knew that the numeral 4 was near the beginning of the number list, but may not have known that it occupied the second ordinal position).

In Chapter 4, two monkeys learned to use a changing Arabic numeral cue (2 through 4) to anticipate when a nonreinforced trial would occur within a series of reinforced and nonreinforced trials. This positive finding contrasted with the failure of a different pair of monkeys to use a larger range of numerals (2 through 8) to anticipate nonreinforced trials. Instead, they performed faster overall on series with higher target numbers. One possible explanation for this behavior is that the monkeys recognized that during the series with higher target numbers, more reinforced trials occurred before the one nonreinforced trial. This may have provided increased motivation, which in turn led to faster performance times. These results could be used as evidence in favor of the object-file model described by Uller, Carey, Huntley-Fenner, & Klatt (1999) because the monkeys were unable to predict the nonreinforced trial on target series higher than four, but we believe they are a better fit with the accumulator model described by Dehaene (1997) and Gallistel and Gelman (2000). If the monkeys were using an object-file mechanism to store each individual trial in a slot in working memory, they would have performed at random when the slots became full. However, the two monkeys receiving probe trials of the numerals 2 through 8 developed a pattern of performance involving all the target numerals. This finding indicates that those monkeys were representing, at least in approximate form, the numerical value of the target numbers 2 through 8.

Evidence that the monkeys represent number with inexact magnitude representations can also be found in the second experiment of Chapter 2. In this experiment, three out of four monkeys performed at better-than-chance level on a task requiring them to make a number of responses equaling an Arabic numeral. Although the monkeys performed better than would be expected if they only knew the relative magnitudes of the numerals, errorless problems were still in the minority. On most trials, the monkeys touched a target more times than was represented by the numeral. Thus, it seems unlikely that the monkeys were enumerating responses toward some exact and absolute quantity (e.g., moving to the 3 exactly three times). Instead, these data suggest that the monkeys were responding to the approximate quantity represented by each numeral.

The same conclusion can be drawn from the experiment in Chapter 5. In this experiment, the monkeys learned to enumerate their own sequential responses and to match the number of responses with the corresponding Arabic numeral or visual quantity. Although the monkeys were able to perform this task, there is evidence that they were using an approximate and variable representation of the number of maze trials. Accuracy increased as a function of the distance between the number of maze trials and the numerosity presented during the discrimination. Accuracy also decreased as a function of the ratio of the smaller numerosity to the larger numerosity used in each series, as predicted by Weber's law. A distance effect and adherence to Weber's law would occur if the monkeys' numerical representations were composed of inexact magnitudes (Dehaene, 1997; Gallistel & Gelman, 2000).

Collectively, these experiments demonstrate that the monkeys do not have a human-like understanding of numerals. Humans use number words and symbols to move beyond the realm of approximation and communicate the precise numerical values required for formal mathematics. To an adult human, the numeral 4 symbolizes exactly four items or actions, which is a quantity that is precisely one unit less than five and one unit more than three. In contrast, the monkeys behaved as if the representations underlying the Arabic numerals were fuzzy approximations of true set size rather than precise quantities.

Monkeys have an inherent disadvantage compared to humans in that they cannot use number words and symbols to communicate precise quantities and perform formal mathematical operations. However, animals in the wild typically confront situations where relative knowledge is sufficient. For example, it is important for an animal to know whether its allies outweigh its foes before engaging in a conflict, but it is not necessary to know the exact number of friends or foes. Thus, in most situations an approximate representation of number provides the information necessary for efficient decision-making.

Another area in which the monkeys may not possess a human-like understanding of number is their ability to generalize numerical knowledge from one task to another. Adult humans have a generalized concept of number that allows them to abstract number across different contexts, even when the numerical stimuli differ in perceptual features and modality. For example, adults understand that three apples, three flashes of lightning, and three trips to the grocery store can all be classified in the same numerical category, even though these items and events occur in different contexts and modalities and do not resemble each other perceptually. The question is whether the nonverbal, less numerically experienced rhesus monkeys in our laboratory possess similar representations of number. Thus, the last step to determining the nature of the monkeys' concept of number was to determine the generality of their representations by testing their ability to abstract number across different presentation modes.

The potential difference in numerical generalization abilities between our monkeys and adult humans became apparent after conducting the ordinal experiment in Chapter 3. The monkeys in that experiment both had experience comparing Arabic numerals (Harris, Gulledge, Beran, & Washburn, 2008), altering their behavior based on a changing target numeral (Harris & Washburn, 2005), and matching an Arabic numeral to the corresponding number of sequential behaviors (Harris, Washburn, Beran, & Sevcik, 2007), but showed no advantage when learning lists of Arabic numerals versus lists of novel colors and signs. The study in Chapter 3 was conducted approximately six months after the monkeys were last exposed to Arabic numerals so it is possible that numerical learning was lost between studies due to time and interference from intervening experimental tasks. This possibility would be surprising, however, in light of the evidence that chimpanzees can retain the values of Arabic numerals for an interval of over three years (Beran, 2004). It is also possible that the type of numerical knowledge learned in the previous tasks was not applicable to the ordinal numerical task. Thus, it was unclear whether the monkeys could transfer numerical knowledge from one context to another.

Findings from the training phase of Chapter 4 also provided no evidence of numerical transfer. In this study, two Arabic numeral-experienced monkeys learned to use a changing target numeral to predict when a nonreinforced trial would occur, but they both performed several hundred series before the pattern emerged. In contrast, rats in similar studies developed the pattern after performing less than 50 series (Burns, Goettl, & Burt, 1995; Capaldi & Miller, 1988). The monkeys may have required more training because in previous tasks a nonreinforced trial signaled an incorrect response. Therefore, the monkeys had to overcome the prior meaning of a nonreinforced trial before learning to predict when it would occur. Another explanation is that the monkeys required extensive training because they were unable to generalize their previous numerical experience to the current task.

In the testing phase of the Chapter 4 study, we introduced probe trials involving a range of Arabic numerals and hypothesized that the monkeys' prior knowledge of Arabic numerals would allow for spontaneous transfer from one Arabic numeral to another during this sequential task. For the monkeys receiving probe trials of the numerals 2 and 4, the first ten probe trials of each novel numeral provided some evidence that immediate generalization to the new numerals occurred. The average performance time for those trials was greater for the last nonreinforced trial in each novel series than the reinforced trial before it. However, the monkeys receiving probe trials of the numerals 2 through 8 failed to generalize to the new numerals.

All of these instances in which the monkeys failed to demonstrate transfer of numerical knowledge from one task to another could be explained by the time elapsed between studies or the fact that the tasks utilized in the different studies had widely varying procedures and goals. Another explanation for the lack of positive findings is that the monkeys form numerical representations that are strongly tied to context and not easily abstracted across different tasks, modalities, and perceptual features. In other words, the numerals are functioning as part of a specific stimulus-response-reward association which allows the monkeys to generalize a numerical rule to novel numerals and combinations of numerals within a task, but does not allow them to abstract that numerical value across different contexts. The study in Chapter 6, in which the monkeys were tested on their ability to transfer numerical knowledge from a task involving sequentially presented stimuli to a very similar task involving simultaneously presented stimuli, was designed to investigate these two competing proposals.

The results of this study revealed that four out of the eight monkeys showed no evidence of transfer from the sequential training task to nonreinforced simultaneous probe trials in Experiment 1. In contrast, three of the monkeys performed at high levels on the simultaneous probe trials during the beginning of testing, but were unable to sustain performance across all 400 probe trials. This suggests that they may have initially been using the reward contingencies learned during the sequential task to perform the simultaneous task, but lost motivation after receiving a large number of nonreinforced trials. The eighth monkey exhibited high levels of accuracy on probe trials throughout testing, which indicated that he was able to generalize his learned responses for sequentially presented circles to the same number of simultaneously presented circles.

In the second experiment, a group of monkeys receiving consistent reinforcement regardless of presentation method (sequential or simultaneous) required approximately the same number of trials to reach the accuracy criterion as a group of monkeys receiving reversed reward contingencies, which provides no evidence of numerical transfer. Overall, these experiments indicate that some rhesus monkeys have an abstract concept of number that reaches across presentation mode, but that concept is not automatically activated in all numerical situations. The fact that the monkeys typically work on several different tasks each day might have interfered with any natural tendency to transfer knowledge from one task to another, regardless of how similar the tasks appeared. It is also possible that monkeys, unlike humans, do not have a true concept of number that allows them to abstract number across different contexts, and therefore the ability is only exhibited by some monkeys under limited conditions.

Together, these studies demonstrate that monkeys have a numerical concept that allows them to perform a wide range of numerical tasks. Although the monkeys have had extensive numerical experience over the course of their lives, they are not a part of the same number-rich culture as their human relatives, who must represent and use number to interact with their world on a daily basis. The success of these monkeys and other nonhuman animals on numerical tasks demonstrates that some numerical capacity is not an entirely cultural construction limited to the human species. Instead, humans and nonverbal animals share a basic system for representing numbers as continuous magnitudes.

Although the numerical abilities of humans and nonhuman animals share an evolutionary past, this does not mean that we should think of animal numerical abilities simply as lesser versions of human numerical abilities. In the field of numerical cognition, as with other fields of cognition, we should be cautious about directly equating nonhuman animal abilities to the abilities of humans at different ages because the different species may have unique capabilities and mental representations. For example, rhesus monkeys and preverbal human infants exhibit a similar ability to differentiate between small quantities of items (e.g., Hauser & Carey, 2003; Starkey & Cooper, 1980; Washburn & Rumbaugh, 1991; Wynn, 1992), but they appear to differ in their ability to abstract number across different contexts. Several studies have shown that infants with no formal number training have a spontaneous ability to abstract number across different presentation methods (Jordan, Brannon, & Gallistel, 2006; Starkey, Spelke, & Gelman, 1983, 1990), whereas the majority of numerically-sophisticated monkeys in our laboratory had difficulty abstracting number from one task to another. The disparity between infants and monkeys could be a function of the different methodologies used in the studies, but it could also represent a fundamental difference between the two species.

Another critical difference is that monkeys trained with Arabic numerals understand the symbolic meaning of the numerals, but there is no evidence that human infants can use symbols to represent quantities. The fact that monkeys use Arabic numerals to perform numerical tasks, however, does not mean that they perceive numerals in the same way as adult humans. Whereas humans use numerals to symbolize precise quantities, we have seen overwhelming evidence that nonhuman animals perform numerical tasks using imprecise magnitudes representations. Thus, it appears that animals map Arabic numerals onto their inexact quantification system rather than using the numerals to develop precise representations. Although studying this inexact system can help us understand the roots of human mathematical abilities, it can also help us understand how monkeys and other animals think and interact with their environments in ways that are unique to them.

Given the results of these studies and findings from the rest of the literature, what do we now know about animal numerical cognition? First and foremost, we know that a wide range of species can perform simple numerical judgments, even when confounding factors such as time and area are taken into account. This knowledge has allowed the field of animal cognition to move well beyond the shadow of Clever Hans. Instead of focusing on whether or not animals *can* use number, today's research focuses more on *how* animals use number, including the nature of their numerical concepts and the underlying mental mechanisms.

The area of research involving experimentally naïve animals has provided a wealth of evidence that the numerical sensitivity exhibited by animals is present without training. However, number is often confounded with other variables in the natural environment so researchers continue to debate whether number is a highly salient cue, or one that is used as a last resort when all other cues fail to provide reliable information. It also remains unclear whether studies involving experimentally naïve animals and small quantities of food items activate the same mental mechanisms as laboratory studies that involve extensive training.

The majority of laboratory studies have shown that animals demonstrate an understanding of both the relative numerical properties of numbers and their approximate cardinal values. These studies also have shown that animals represent cardinal number using an analog magnitude mechanism, but the exact nature of this mechanism remains under debate, with some researchers arguing for a logarithmic mental number line and others for a linear mental number line. In addition, recent research has provided evidence that nonhuman primates understand the symbolic nature of Arabic numerals and other abstract symbols associated with numerical quantities, and can use these symbols to perform a variety of tasks.

Another important finding from the field of animal numerical cognition is that animals can perform a wide variety of numerical tasks including visual quantity judgments, sequential quantity judgments, and auditory numerical discriminations. Although there is some evidence that animals have an abstract, amodal representation that can be transferred across tasks, the contexts under which animals exhibit or do not exhibit this transfer ability have yet to be specified.

Not only do these studies provide greater perspective on the current state of the field, but they also provide insight into potential future directions for animal numerical research that would enhance our overall knowledge and accelerate progress in the field. It is clear that animal numerical cognition is a highly interdisciplinary area of research that includes investigators from a variety of backgrounds including comparative psychology, developmental psychology, cognitive science, neuroscience, anthropology, and biology, who research a wide range of animal species. Although research backgrounds and subject species vary widely throughout the field, individual researchers tend to narrowly focus on one research paradigm when conducting numerical studies. For instance, one researcher might rely on tasks involving visually presented quantities whereas another relies on tasks involving sequentially completed movements and yet another focuses on ordinal sequencing paradigms. These divisions within the field hinder the ability of researchers to integrate findings from different facets of numerical research.

The studies in this dissertation demonstrated how findings from simultaneous, sequential, and ordinal studies conducted with the same monkey population could be combined to increase our knowledge of larger trends in numerical cognition. For example, the lack of numerical transfer from one type of study to another prompted the formal study of numerical abstraction in Chapter 6. This trend would not have been evident if we had not presented the same population of monkeys with tasks involving a wide variety of paradigms. In addition, the fact that widely varying tasks each produced results that conformed to the predictions of the analog magnitude model provides evidence that this may be an overarching mechanism involved in all numerical judgments.

The flexibility of the analog magnitude model, along with the predictions it makes regarding distance and magnitude effects, is one of the reasons it has been widely respected since it was first proposed by Meck and Church (1983). As opposed to other models, such as the subitizing model which describes numerical processing mainly in terms of visual perception, the accumulator mechanism easily accommodates data from both simultaneous and sequential tasks. According to the analog magnitude model, organisms possess an internal pacemaker that emits a stream of pulses at a steady rate into a mental accumulator. In a simultaneous task, the gate to the internal pacemaker is opened and closed after each individual object in the array is enumerated. Thus, the level of the accumulator is correlated with the number of objects in the array. In a sequential task, the gate is opened and closed after each object is encountered, which also produces an accumulator level that is correlated with the number of objects in the set. In an ordinal or relative numerous task in which the subject must compare or order two or more quantities, without necessarily knowing the absolute value of either quantity, subjects need only compare the two mental magnitudes generated by the accumulator. This judgment becomes less accurate as the numerical distance between the quantities decreases and the value of the quantities increases because the memory for the magnitudes is imperfect and defined by scalar variability.

Although there is a great deal of evidence that animals use analog magnitude representations to perform numerical tasks, questions remain regarding the exact form of those representations. In other words, when a monkey thinks about number, what exactly does that mean? Does the monkey visualize a physical quantity with area and density, a point along a mental number line, or something completely different? Do their numerical representations always take the same form, or do they differ based on the context? Varying forms of numerical representations could explain the negative findings in the numerical abstraction study. It is possible that the monkeys were not able to generalize numerical learning from the sequential to the simultaneous task because the two tasks activated different analog representations. In addition to questions about the exact form of the analog representations, there are also questions regarding the potential for other forms of representation.

According to the triple-code model proposed by Dehaene (1992), adult humans represent numbers in a visual, verbal, or analog magnitude code depending on the task. These different types of numerical representation, which are processed by different areas of the brain, are supported by specific comprehension and production mechanisms and connected by pathways that allow translation from one type of representation to another. In this view, the visual representations used to process Arabic numerals in multi-digit operations consist of strings of Arabic digits. The verbal representations used to process spoken and written number words, as well as to count and solve simple addition and multiplication problems, consist of sequences of number words. Finally, analog representations used to perform approximate calculations and relative numerousness judgments consist of mental continuums that are compressed near the larger numbers.

It is clear that nonverbal animals do not have verbal representations they can use to count and perform simple calculations, but things are less clear regarding visual representations. Although animals trained with Arabic numerals can learn decision rules that allow them to assign a specific symbol to a range of accumulator values, there is currently no evidence that they have visual representations, similar to those proposed for humans, which would allow them to perform precise numerical judgments and calculations. It is interesting to note, however, that humans learn the meaning of visual symbols such as Arabic numerals through constant training and experience at school and at home. In fact, several developmental researchers have stated that a major part of learning to count involves learning to map back and forth from magnitude representations to numerals (Gallistel & Gelman, 1992; Whalen et al., 1999). If humans can learn to represent number precisely using symbols over the course of development, then it is possible that rhesus monkeys, with the right training and experience, may also be able to make this conceptual leap. Thus, it is possible that monkeys, like humans, may be able to build on the foundations of the analog code and learn to use other forms of numerical representation.

In conclusion, the studies in this dissertation further defined the extent of numerical ability in rhesus monkeys and the nature of their numerical concept. It is clear that numerical ability is not based on low-level associative or perceptual processes and that stimulus control by the numerousness aspect of stimuli is not the same as control by other dimensions. Instead, rhesus monkeys have conceptual numerical knowledge that guides behavior in a variety of number-related contexts, and in some cases, allowed the monkeys to generalize a response rule across presentation mode. In addition, the monkeys recognized that the Arabic numeral cues used in the tasks symbolized numerical quantities with ordinal and approximate cardinal value. Collectively, these studies provide evidence that rhesus monkeys view Arabic numerals as more than conditioned stimuli with specific response-reward histories, but that numerals do not have the same precise symbolic meaning as they typically do for humans.

References

Anderson, U. S., Stoinski, T. S., Bloomsmith, M. A., Marr, M. J., Smith, A. D., & Maple, T. L. (2005). Relative numerousness judgment and summation in young and old western lowland gorillas. *Journal of Comparative Psychology, 119,* 285-295.

Beran, M. J. (2004). Long-term retention of the differential values of Arabic numerals by chimpanzees (*Pan troglodytes*). *Animal Cognition, 7,* 86-92.

Boysen, S. T., & Berntson, G. G. (1995). Responses to quantity: Perceptual versus cognitive mechanisms in chimpanzees (*Pan troglodytes*). *Journal of Experimental Psychology: Animal Behavior Processes, 21*, 82-86.

Brannon, E. M., & Terrace, H. S. (1998). Ordering of the numerosities 1–9 by monkeys. *Science, 282*, 746–749.

Brannon, E. M., & Terrace, H. S. (2000). Representation of the numerosities 1-9 by rhesus macaques (*Macaca mulatta*). *Journal of Experimental Psychology: Animal Behavior Processes, 26,* 31–49.

Burns, R. A., Goettl, M. E., & Burt, S. T. (1995). Numerical discrimination with arrhythmic serial presentations. *Psychological Record, 45,* 95-104.

Capaldi, E. J., & Miller, D. J. (1988). Counting in rats: Its functional significance and the independent cognitive processes that constitute it. *Journal of Experimental Psychology: Animal Behavior Processes, 14,* 3-17.

D’Amato, M. R., & Colombo, M. (1988). Representation of serial order in monkeys (*Cebus apella*). *Journal of Experimental Psychology: Animal Behavior Processes, 14,* 131–139.

Davis, H., & Pérusse, R. (1988). Numerical competence in animals: Definitional issues, current evidence, and a new research agenda. *Behavioral and Brain Sciences, 11,* 561-615.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition, 44,* 1-42.

Dehaene, S. (1997). *The number sense: How the mind creates mathematics.* New York: Oxford University Press.

Flombaum, J. I., Junge, J. A., & Hauser, M. D. (2005). Rhesus monkeys (*Macaca mulatta*) spontaneously compute addition operations over large numbers. *Cognition, 97,* 315-325.

Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition, 44,* 43-74.

Gallistel, C. R., & Gelman, R. (2000). Non-verbal numerical cognition: From reals to integers. *Trends in Cognitive Sciences, 4,* 59-65.

Harris, E. H., Gulledge, J. P., Beran, M. J., & Washburn, D. A. (2008). *What do Arabic numerals mean to macaques?* Manuscript submitted for publication.

Harris, E. H., & Washburn, D. A. (2005). Macaques' (*Macaca mulatta*) use of numerical cues in maze trials. *Animal Cognition, 8,* 190-199.

Harris, E. H., Washburn, D. A., Beran, M. J., & Sevcik, R. A. (2007). Rhesus monkeys (*Macaca mulatta*) select Arabic numerals or visual quantities corresponding to a number of sequentially completed maze trials. *Learning and Behavior, 35,* 53-59.

Hauser, M. D., & Carey, S. (2003). Spontaneous representations of small numbers of objects by rhesus macaques: Examinations of content and format. *Cognitive Psychology, 47,* 367-401.

Hauser, M. D., Carey, S., & Hauser, L. B. (2000). Spontaneous number representation in semi-free-ranging rhesus monkeys. *Proceedings of the Royal Society of London B, 267,* 829-833.

Jordan, K. E., Brannon, E. M., & Gallistel, C. R. (2006). The multisensory representation of number in infancy. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, *103,* 3486-3489.

Judge, P. G., Evans, T. A., & Vyas, D. K. (2005). Ordinal representation of numeric quantities by brown capuchin monkeys (*Cebus apella*). *Journal of Experimental Psychology: Animal Behavior Processes, 31,* 79-94.

Meck, W. H., & Church, R. M. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behavior Processes, 9,* 320-324.

Menzel, E. W., Jr. (1960). Selection of food size in the chimpanzee and comparison with human judgments. *Science, 131,* 1527-1528.

Nieder, A., & Miller, E. K. (2004). Analog numerical representations in rhesus monkeys: Evidence for parallel processing. *Journal of Cognitive Neuroscience*, *16,* 889-901.

Santos, L. R., Barnes, J. L., & Mahajan, N. (2005). Expectations about numerical events in four lemur species (*Eulemur fulvus, Eulemur mongoz, Lemur catta and Varecia rubra*). *Animal Cognition, 8,* 253-262.

Smith, B. R., Piel, A. K., & Candland, D. K. (2003). Numerity of a socially housed hamadryas baboon (*Papio hamadryas*) and a socially housed squirrel monkey (*Saimiri sciureus*). *Journal of Comparative Psychology, 117,* 217-225.

Starkey, P., & Cooper, R. G. (1980). Perception of numbers by human infants. *Science, 210,* 1033-1035.

Starkey, P., Spelke, E. S., & Gelman, R. (1983). Detection of intermodal numerical correspondences by human infants. *Science, 222,* 179-181.

Starkey, P., Spelke, E. S., & Gelman, R. (1990). Numerical abstraction by human infants. *Cognition, 36,* 97-127.

Terrace, H. S., Son, L. K., & Brannon, E. M. (2003). Serial expertise of rhesus macaques. *Psychological Science, 14,* 66–73.

Uller, C., Carey, S., Huntley-Fenner, G., & Klatt, L. (1999). What representations might underlie infant numerical knowledge? *Cognitive Development, 14,* 1-36.

Uller, C., Hauser, M., & Carey, S. (2001). Spontaneous representation of number in cotton-top tamarins (*Saguinus oedipus*). *Journal of Comparative Psychology, 115,* 248-257.

Washburn, D. A., & Rumbaugh, D. M. (1991). Ordinal judgments of numerical symbols by macaques (*Macaca mulatta*). *Psychological Science, 2,* 190–193.

Whalen, J., Gallistel, C. R., & Gelman, R. (1999). Nonverbal counting in humans: The psychophysics of number representation. *Psychological Science, 10,* 130-137.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature, 358,* 749-750.

Xia, L., Siemann, M., & Delius, J. D. (2000). Matching of numerical symbols with number of responses by pigeons. *Animal Cognition, 3,* 35-43.