

COMPARATIVE COGNITION & BEHAVIOR REVIEWS

2010

Volume 5, pp 23-58

The Comparative Psychology of Serially Organized Behavior

Herbert Terrace

Columbia University & the New York State Psychiatric Institute

The study of serially organized behavior has benefited from a new paradigm for training sequences, from new technology for presenting multiple images in varied spatial positions and from new concepts for describing serially organized behavior. The new paradigm is the simultaneous chaining paradigm, one that presents all list items simultaneously, in a new configuration on each trial. Because there are no external cues to guide the execution of the required sequence, subjects must form a representation of the sequence and update it while moving from item to item. Experiments in which humans and monkeys were trained to learn sequences composed of arbitrary items showed that subjects acquired knowledge of the ordinal position of each item, and its relationship with other items from a list, without any requirement to do so. Symbolic distance and magnitude functions, that were obtained from both monkeys and humans, who were trained to execute arbitrary and numerical lists, provide strong evidence of an underlying ordinal knowledge, at both the behavioral and the neural level.

Serial learning is one of the oldest and most widely studied phenomena of experimental psychology; as well it deserves to be. As compared to individual responses, serially organized action is the norm in everyday behavior and it is also fundamental for the mastery of skills at all levels of complexity; skills as simple as knowing how to get from point A to point B and as complex as knowing how to speak and comprehend language.

Until recently, the only available models of serially organized behavior in animals were based on associations between successive list items. Like the model introduced by Ebbinghaus (1964), in his research on human memory, those models, assumed that sequences were simply chains of associations between adjacent items, and nothing more. Pavlov's discovery of the conditioned reflex (Pavlov, 1927) inspired a research program on properties of learned asso-

ciations, a program that underwent considerable expansion with Skinner's discovery of operant conditioning (Skinner 1938). That program was remarkably successful. It is no exaggeration to say that the theories of classical and operant conditioning of individual responses that emerged from that research is one of the major achievements of 20th century psychology. Common to all such theories is the assumption that learning, in both animals and humans, can be characterized as the establishment of stimulus-response (S-R), response-stimulus (R-S), or stimulus-stimulus (S-S) associations, or some combination thereof. Although the particulars of associations differ in competing theories, they all focus exclusively on individual responses

During the latter part of the 20th century, chaining theory was supplemented with a cognitive view of serial learning that was based on knowledge of the ordinal position of each item and their spatial representation. Instead of focusing exclusively on how one list item is connected to the next, a cognitive model seeks to explain knowledge of relationships between non-adjacent items. When, for example you run to catch a ball, you don't take one step at a time. Instead, you relate the path of your run with the trajectory of the ball so that you both end up at the same place at the same time. Indeed, chaining theory could not explain how an outfielder

Herbert Terrace, Department of Psychology, Columbia University & the New York State Psychiatric Institute.

This research was supported by Grant R0 MH081153-1A1 from the National Institute of Mental Health.

Correspondence concerning this article should be addressed to Herbert S. Terrace, Columbia University, 406 Schermerhorn Hall, 1190 Amsterdam Ave., New York, NY 10027, E-mail: terrace@columbia.edu

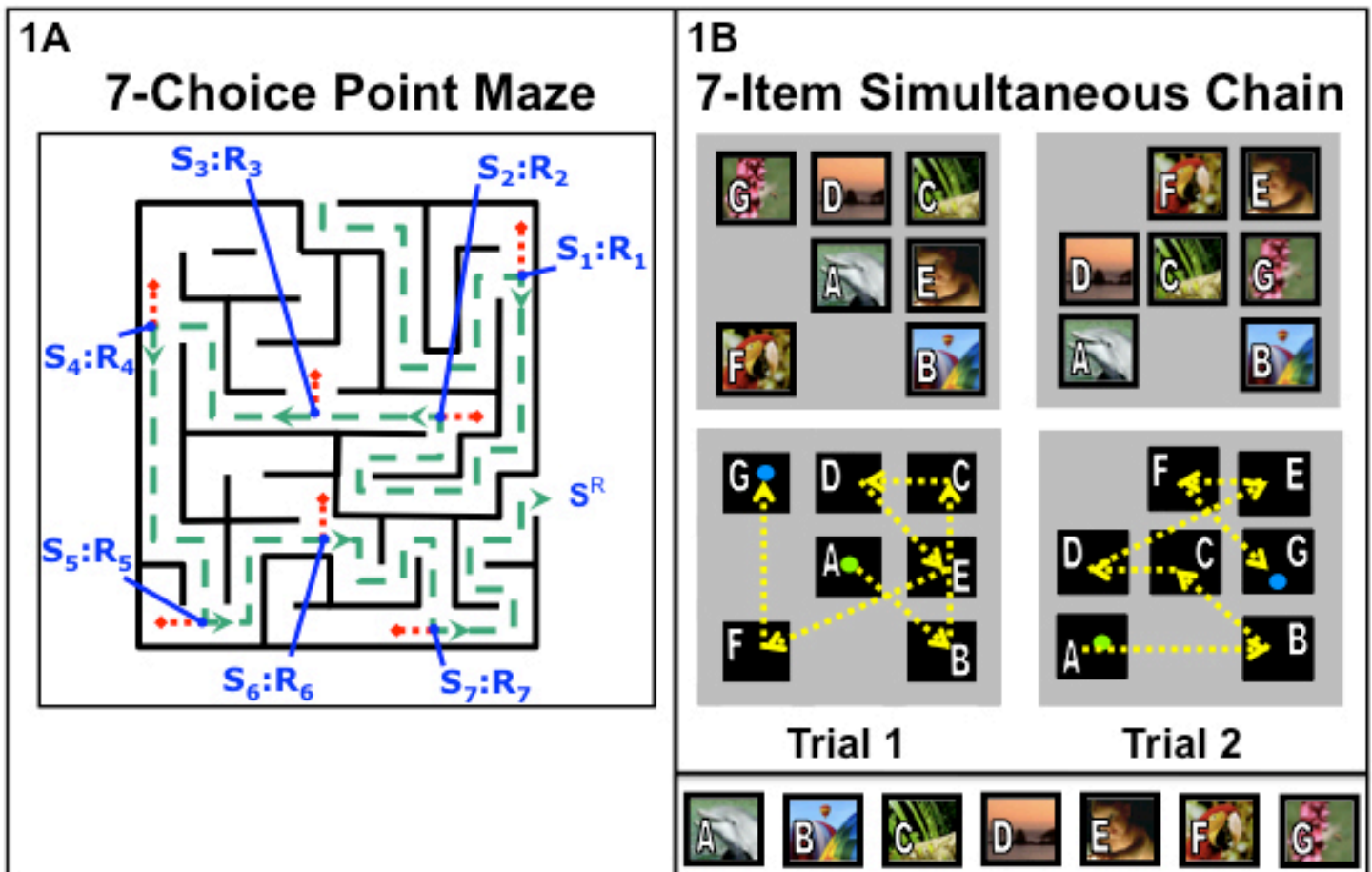


Figure 1a. *What a subject learns in a maze with n choice points.* According to chaining theory, a subject learns n discrete stimulus–response (S – R) associations while learning to navigate a maze. In the maze shown here, learning 7 associations allows the subject to arrive at the goal box by executing the following successive chain: $S_1:R_1 \rightarrow S_2:R_2 \rightarrow S_3 \dots S_7:R_7 \rightarrow S^R$. Crucially, this knowledge also insures that (1) having responded correctly to S_n , that stimulus disappears and (2) encounter S_{n+1} , and only S_{n+1} , will be encountered at the next choice point. Thus, when the subject is at, say, S_3 , cues from the other choice points cannot compete for attention.

Figure 1b. *Simultaneous chaining paradigm.* Each panel depicts a trial during training on a given simultaneous chain. The configuration of list items, typically photographs, varies randomly from trial to trial. The photographs at the bottom of the figure show, for the benefit of the reader, the correct order in which a subject must respond in order to obtain a reward. The yellow arrows in the bottom portion of each panel depicts the route that a subject must follow with respect to the items shown in the top panel in order to receive a reward. Barring a series of lucky guesses that by chance would happen only once in 1 in 5040 guesses, a simultaneous chain must be learned by trial and error. A reward is presented only after the subject has responded, to each item, in the correct order. An error at any point of the sequence ends the trial. See text for additional details.

could catch a ball that is thrown at different distances and at different angles because it would be impossible to practice the individual steps required for each catch. Likewise, when you learn to sing, you sing melodies rather than individual notes. These and similar examples have convinced many psychologists that a sequence of behavior is more than the sum of isolated and independent associations.

In one of the most influential articles in modern psychology, Lashley (1951) discussed what he regarded as the fatal flaw of chaining theory. He argued that chaining theory could

not account for knowledge of relationships between non-adjacent items in serially organized behavior. He presented two important examples to support his argument. First, he noted that all languages assume knowledge of relationships between words from different parts of a sentence. For example, if someone told you that the girl, at the ice cream store, wearing the blue uniform, won the race, you would understand that it was the girl, and not the blue uniform, that won the race. Lashley also cited various human skills that could not be characterized as sequences of chained S – R units because the inter-response times between successive responses

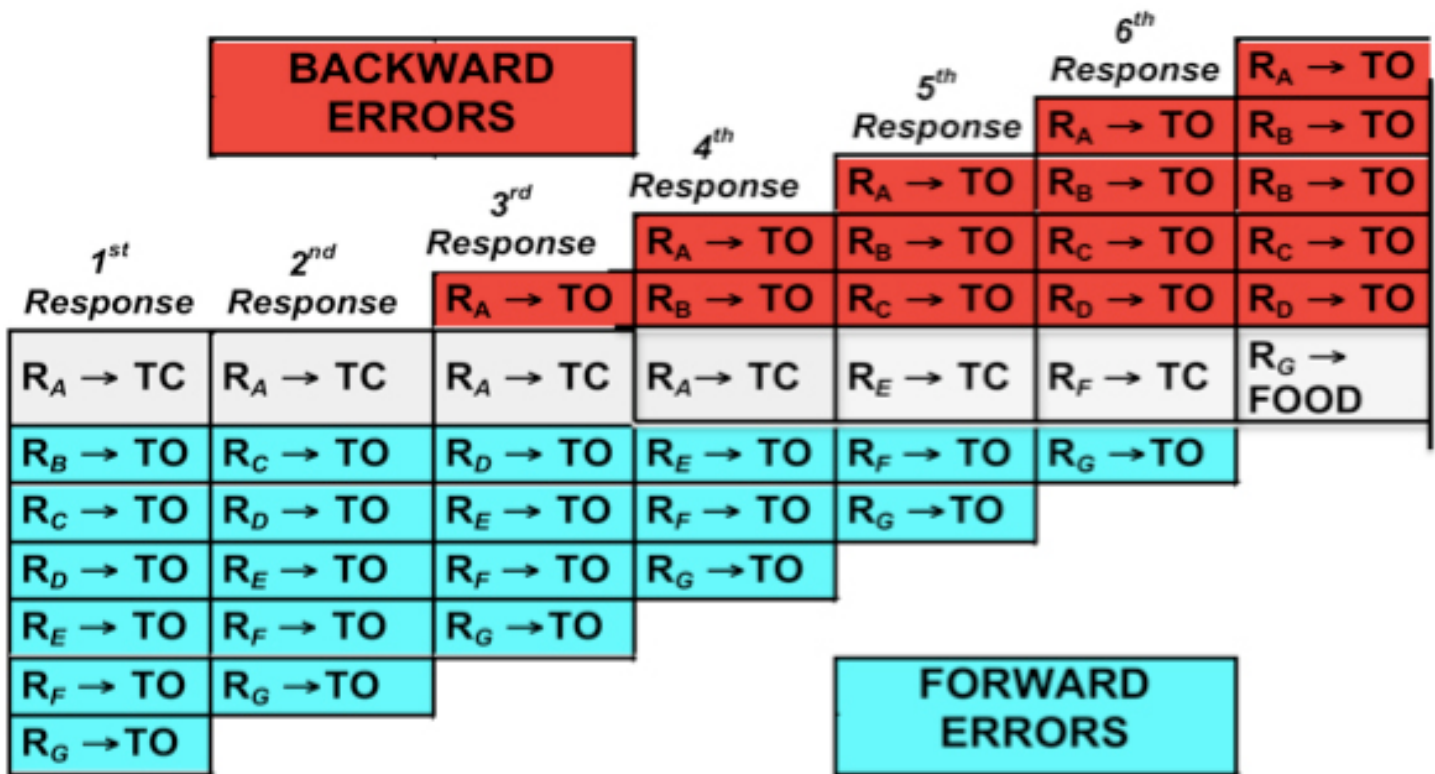


Figure 2. The 21 possible forward and the 15 possible backward errors that can occur during the acquisition of new 7-item simultaneous chain. Because subjects make very few backward errors after mastering a few simultaneous chains, our focus will be on two types of forward error. Logical (or search) errors eliminate various items at particular positions of the sequence. Perseverative errors are responses to items whose position was defined by a prior logical error. If the monkey selects A with the first guess, he will have made 0 logical errors; with the 2nd guess, 1 logical error; with the 3rd guess, 2 logical errors, and so on. The minimum number of logical errors is the sum of the number of possible logical errors that a subject can make while learning the 7-item chain. For example, guessing A after 2 search errors could result from any sequence in which the monkey responds to two different items other than A (B & C, C & D, B & E, etc.). In general, the minimum number of search errors needed to determine A is $6 + 5 + 4 + 3 + 2 + 1 + 0 = 21$. The expected number of logical errors needed to determine A is $21 * 1/7 = 3$; to determine B, $15 * 1/6 = 2.5$; to determine C, $10 * 1/5 = 2$, and so on. The value of the expected number of logical errors at each position decreases linearly, 0.5 guesses at each position, until it reaches a value of 0 at item F.

were often shorter than the time it would take for feedback from one response to trigger the next (e.g. typing or playing a musical instrument).

Because Lashley's arguments were based on examples of human behavior, his critique had less influence on animal than on human cognition. Indeed, it has been suggested that Lashley's arguments do not apply to animals because there is no evidence that they engage in learned behavior that approaches the complexity of human skills, and because animal communication is simpler and less flexible than human language (Lewandowsky 1994). That view is no longer tenable. As we will see, recent advances in our understanding of serially organized behavior in animals have shown that Lashley's criticisms of chaining theory apply with the same force to animal behavior as they do to human behavior.

During the last 25 years, new technology made it possible

to develop new paradigms for studying serial learning in animals. Experiments based on those paradigms have shown, for the first time, that non-human primates share with humans the ability to memorize lengthy sequences composed of arbitrary stimuli (Terrace, Son et al. 2003) and to order numerically defined stimuli according to a numerical rule (Brannon and Terrace 1998). They've also shown striking parallels between the representations used by non-human and human primates when learning how to order arbitrary and numerical stimuli. These and other experiments provided convincing evidence that the capacity of non-human primates to acquire ordinal knowledge is considerably greater than one would expect from a non-verbal animal.

Overview of article: The goal of this article is to describe how the simultaneous chaining paradigm, a new paradigm for training monkeys to learn lists has changed our view of

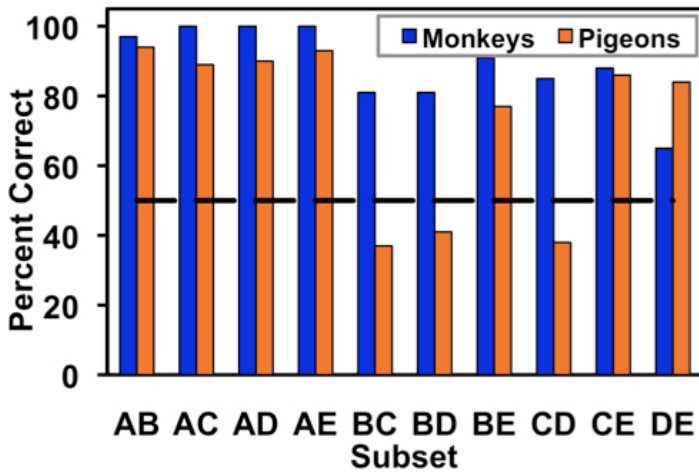


Figure 3. Percentage of correct responses on the pair-wise (2-item) test by four monkeys of the D'Amato and Colombo (1988b) experiment and the 10 pigeons of Terrace's (1987) experiment on each of the 10 test pairs that can be generated from 5-items. Figure from D'Amato, (1991).

the kinds of knowledge that non-verbal primates acquire while mastering a serial task. These lists, which were composed of arbitrary or numerical stimuli that were, respectively, photographs and sets of geometric forms of different size and color, were the longest lists that non-human primates have mastered. As such, they provided an unprecedented opportunity to investigate subjects' knowledge of the relationship between adjacent and non-adjacent list items by testing them on their ability to order two-item subsets that can be derived from each list. These tests have shown that, during the course of learning a list, monkeys acquire knowledge of each item's ordinal position without any verbal ability and without any requirement to do so. Indeed, ordinal knowledge will be a basic theme of this chapter because, more than any other concept, it defines the difference between chaining theory (Hull, 1935, Skinner, 1938) and more modern views of serially organized behavior. As we shall see, it also provides a novel perspective for establishing a comparative psychology of serial learning.

Simultaneous Chains. Much of the research described in this chapter is based on sequences that were trained as simultaneous chains, so named, because all of the list items are presented simultaneously throughout each trial. Examples of a simultaneous chain include entering your password at an ATM, mixing ingredients for a recipe, playing a piano sonata from memory, integrating various maps to plan a trip, and so on. To bring the properties of a simultaneous chain into sharp focus, I will contrast it with the traditional method for training sequences, a method that I have referred to as the successive chaining paradigm.

Maze learning is an example of a successive chain par excellence. When an animal runs through a maze, it encounters

stimuli to which it responds individually and successively to find its way to the goal box (individually in the sense that the animal is not distracted by other choice points that cannot be seen at choice points). Mastering a maze is analogous to finding your way home in a strange town. As a hypothetical example, suppose that driving from the center of town to your hotel requires that you remember what to do at 7 choice points: say, turning left at the post office, right at the bank, right at the gas station, left after crossing a bridge, left at the playground, right at the library, and right after the high school. Your memory of which way to turn at each physically distinct choice point is all you need to know to return to your hotel. In that sense, learning that series of turns requires nothing more than the skills a rat needs to navigate a 7-choice point maze. As can be seen in Figure 1A, responding correctly at the first choice point ensures that you will next encounter the 2nd choice point and only that choice point; responding correctly at the 2nd choice point, ensures that you will next encounter the 3rd choice point, and only that choice point, and so on, until you return to your hotel. What you learn can be characterized as a sequence of S-R associations,

$S1:R1 \rightarrow S2:R2 \rightarrow S3:R3 \rightarrow S4:R4 \rightarrow S5:R5 \rightarrow S6:R6 \rightarrow S7:R7 \rightarrow SR$ (Skinner, 1938) where S_n is a symbol for a particular choice point and R_n a symbol for the correct response at that choice point. Because each response is cued by distinctive stimuli at each choice point, there is no reason to form a representation of the maze Hunter (1920), Terrace (1984).

As shown in Figure 1A, when a subject makes a correct response at one choice point, the subject gains access to the next one. As a result, cues from the previous choice point cannot compete for attention with cues from the current choice point. Along with other behavioral psychologists, Skinner (1938) argued that all instances of serially organized behavior were reducible to discrete S-R units that are linked to the next unit by virtue of practice. It mattered not whether the sequence in question is tying a shoelace, reciting a string of nonsense syllables or producing a grammatical sentence.

Cognitive Maps. Tolman and his students performed a series of experiments to test Skinner's model of maze learning. The basic method was to prevent a rat from reaching particular choice points in a maze by blocking their normal route (Tolman, Ritchie et al. 1946; Tolman 1948). The rat's ability to find alternative paths immediately after a choice point was blocked led Tolman to question the adequacy of the successive chaining model. Instead, Tolman concluded that rats form a cognitive map of the maze and that the cognitive map functioned as a spatial representation of the relationships between different choice points. A cognitive map allows us

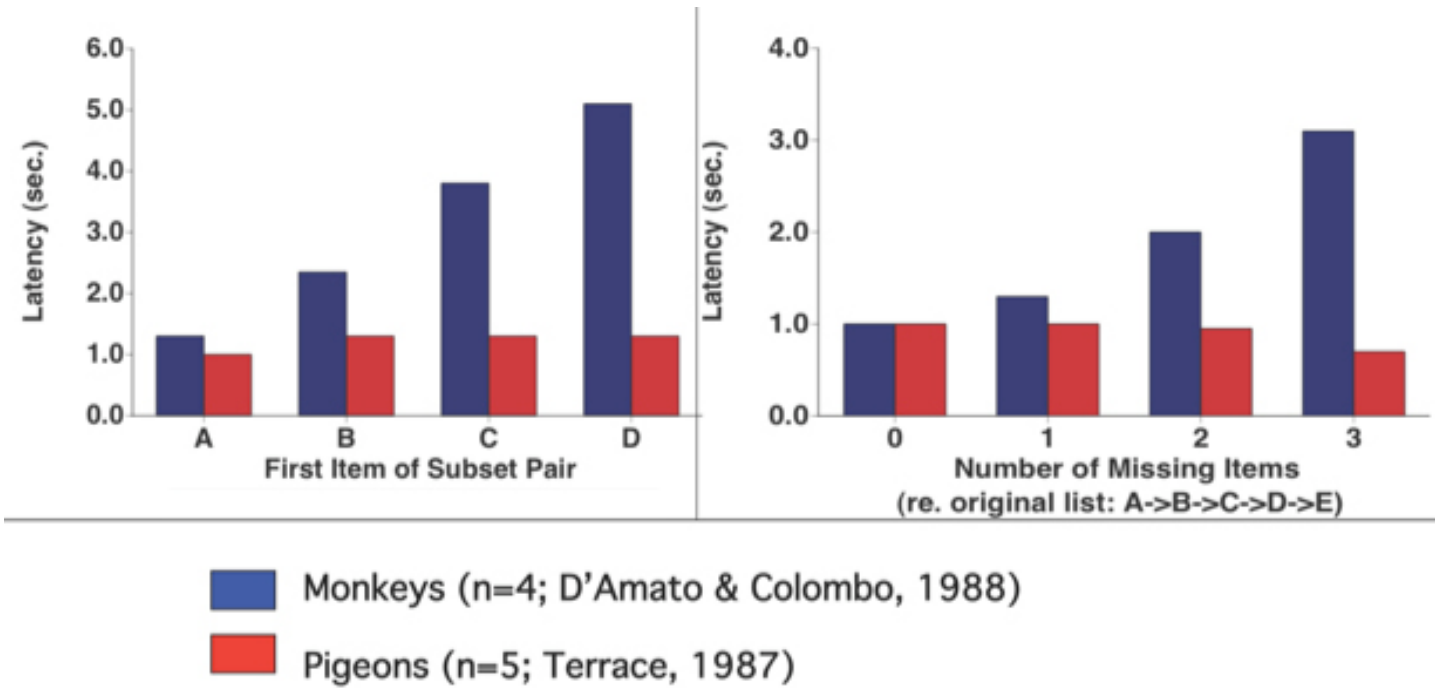


Figure 4. Reaction times (RTs) of monkeys and pigeons following training (on 5-item simultaneous chains) as a function of the physical distance between list items. The blue bars show data from monkeys (D'Amato and Colombo, 1994); the red bars from pigeons (Terrace, 1987).

to construct and accumulate spatially defined images whose function is to enhance recall and learning of information. This type of spatial thinking can also be used in non-spatial tasks by learning mnemonics that locate unrelated list items in a well-known spatial configuration, as for example, in the method of loci (Yates, 1966)

Tolman's insights did not bear fruit until 1969, when his hypothesis about cognitive maps was confirmed in experiments on a rat's ability to navigate a radial arm maze (Olton and Samuelson 1976). As we shall see, the cognitive maps that rats use to navigate a radial arm maze bear a strong resemblance to the representations that a monkey learns while learning a new simultaneous chain.

Lashley's criticisms of association theory underscored the need for a different approach for describing learned sequences that could be scaled up to the types of sequences that are common in human behavior. I designed the simultaneous chain with that goal in mind. The sine qua non of a simultaneous chain is that all list items are displayed throughout each trial and that there is no differential feedback following each response as to the identity of the next item. The only consequence of a correct response is the continuation of the current trial. All errors terminate a trial immediately without providing any information as to the correct response at that point in the sequence. These constraints make it necessary for the subject to construct a representation of the required sequence and to maintain its position within that representation as it advances from item to item. As such, they are

roughly analogous to the constraints that guide our selection of the words when producing a sentence. Those words are simultaneously available in memory and can be chosen in different orders according to particular grammatical rules. This is not to claim that a simultaneous chain is a sentence. It is clearly not. But a simultaneous chain requires the ability to represent an arbitrary sequence, an ability that has, thus far, only been observed in non-human primates. By contrast, a representation is not required for the execution of the successive chains on which rats, pigeons and other non-primates have been trained.

Imagine, for example, what would happen if you forgot your newly-issued 7-item PIN, which in this instance was a series of photographs as shown in Figure 1B. Since the numbers used to compose passwords have no meaning, they are just as arbitrary as arbitrarily chosen photographs, letters or any other arbitrary symbol.

Each photograph covers the number that would normally appear on the number pad. Thus, instead of learning numerical password, say, 9-2-1-5-8-4-7, you are required to learn a sequence of photographs, say, dolphin-hot air balloon-plant-ocean-baby-parrot-flower. Also, as shown in Figure 1B, the spatial location of each photograph changes randomly each time you tried to obtain cash from this ATM. That insures that you can't enter your password by using a fixed motor-pattern, that is, you can't mindlessly "let your fingers do the work". To determine your 7-item PIN, you would have to recall the consequences of any of the 36 types of logical errors

you could make while attempting to produce the required sequence (21 types of forward errors and 15 types of backward errors). Exemplars of each type of error are shown in Figure 2. You would also have to learn the first 6 digits of your password without getting any money from the cash machine. With but one exception, this is precisely the problem that monkeys have to solve when they begin training on a new 7-item list. Instead of cash, they get banana pellets.

As shown in Figure 1B, an important differences between a simultaneous chain and a successive chain is the manner in which a subject encounters successive choice points of a sequence and the consequences of correct and incorrect responses at each choice point. During training on a simultaneous chain, all of the choice points are displayed simultaneously throughout each trial. When, for example, a subject is trained to produce a 7-item simultaneous chain, the stimuli S1S2S3S4S5S6S7 are presented simultaneously at the start of each trial, and they continue to appear until the subject either makes an error, at which point the trial ends, or until it earns a reward, by responding to the 7 stimuli in the sequence, R1 → R2 → R3 → R4 → R5 → R6 → R7. This simultaneous chain would therefore be depicted as follows:

S1 S2 S3 S4 S5 S6 S7: R1 → R2 → R3 → R4 → R5 → R6 → R7 → SR (Terrace, 1984)

Experiments with Pigeons. The simultaneous chaining paradigm was first used in an experiment whose purpose

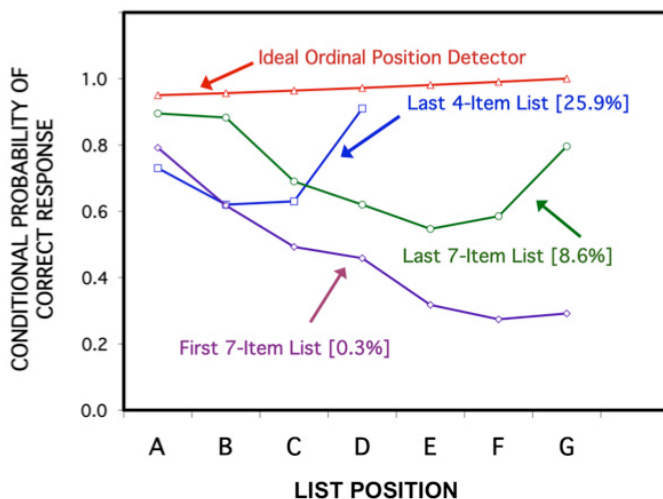


Figure 6. Each function shows the average conditional probability of a correct response at each position during training on the 1st and 4th 7-item lists training and on the last 4-item list (overall list 22). The numbers in brackets to the right of each function show the probability of responding correctly to all of the items on a particular trial (the product of the conditional probabilities of responding correctly to each item).

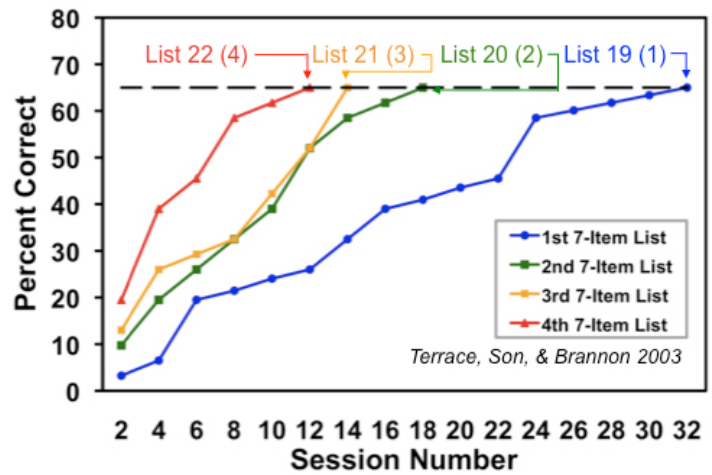


Figure 5. Learning curves for four 7-item lists that are based on the mean accuracy of responding on each 7-item list during even-numbered sessions. The probability of executing a new 7-item list correctly by chance, assuming no backward errors, is $1/5,040$ ($1/7!$). Note that the abscissa represents session (not list number). (Terrace, Son & Brannon, 2003).

was to question claims about the grammatical competence of apes, in particular, the claim that sequences trained by rote were sentences (Straub, Seidenberg et al., 1979). To simulate the conditions used in ape language experiments, Straub et al. trained pigeons to respond to randomly configured arrays of four colors in a particular sequence: red → green → yellow → blue. The probability of guessing the first item correctly is $1/4$. Because repetitive responses to the same items are not considered errors, the probability of guessing B following a correct response to A is $1/3$. Given the generous assumption that subjects don't respond to items to which they've responded previously, the probability of guessing the entire sequence is a mere $.04 = 1/4 \times 1/3 \times 1/2 \times 1$. To avoid the risk of extinction at the start of training, each list was introduced one item at a time. A new item was added to the end of a partial list each time the subject satisfied an accuracy criterion. On a 4-item list, the successive phases of training were A, A → B, A → B → C, and A → B → C → D. This procedure, which I referred to as the successive phase method, was used successfully with monkeys in an experiment conducted by D'Amato & Colombo (1988).

All of the subjects of the Straub, et al. experiment learned to respond correctly to the 4-item list on which they were trained. That result suggested a simpler interpretation of experiments purporting to demonstrate the grammatical ability of apes. Why interpret sequences of plastic chips or lexigrams that apes were trained to produce as anything more complicated than simultaneous chains that were learned by rote whose function was to obtain some specific reward [e.g., Mary → give → Sarah → apple (Premack, 1976) or Please → machine → give → apple (Rumbaugh, 1977)]. As Terrace (1979) noted, there is no evidence that the first 3 symbols of

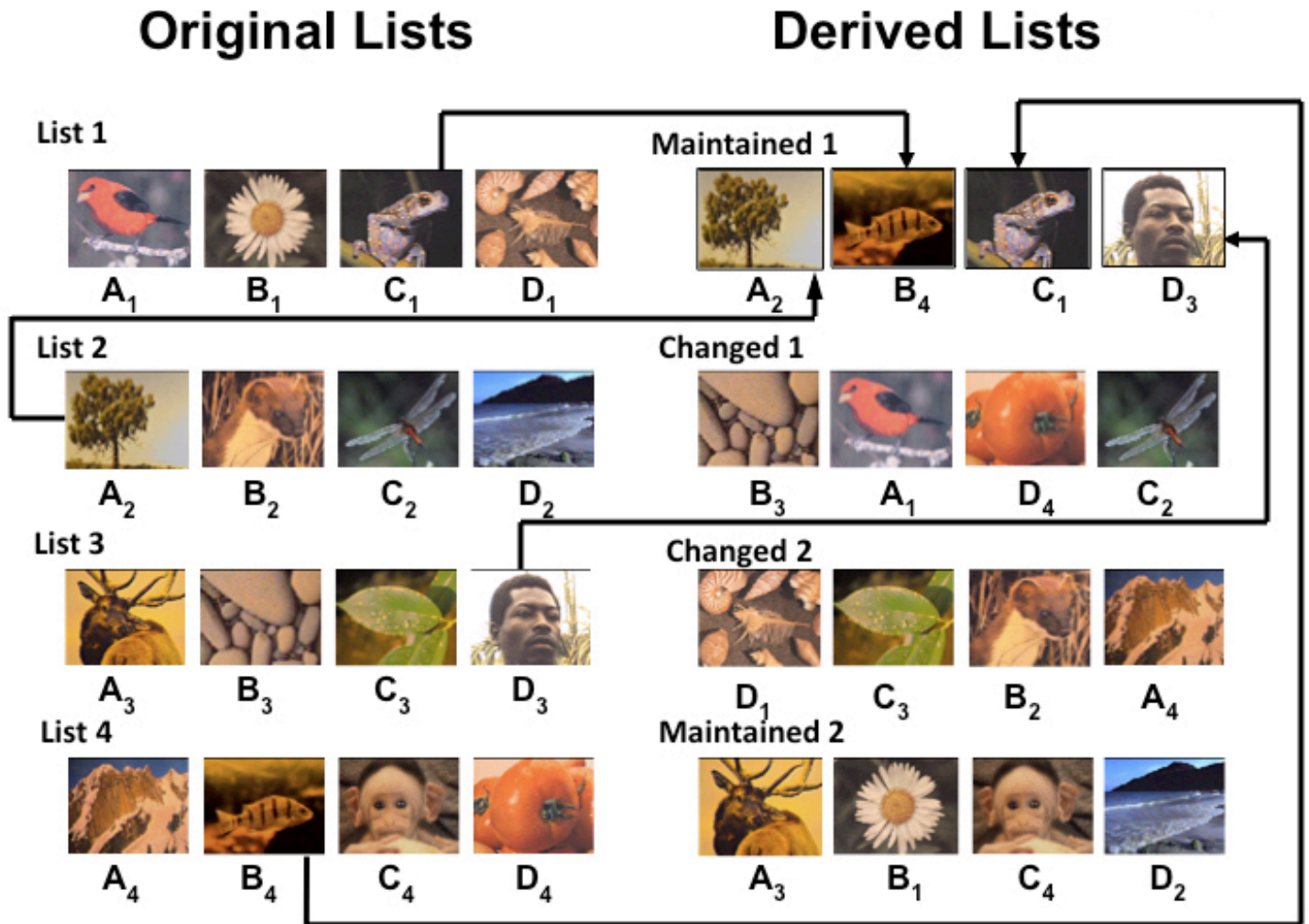


Figure 7. Composition of original and derived lists (Chen, et al, 1997). Maintained and changed lists were presented in an ABBA order.

these sequences had any meaning for the apes in question.

Terrace (1987) extended Straub et al.'s findings by evaluating subjects' knowledge of associations between non-adjacent items on a 5-item list. After subjects satisfied an accuracy criterion, Terrace administered a subset test that consisted of the ten 2-item pairs that could be derived from a 5-item list (AB, AC, AD, AE, BC, BD, BE & CD, CE and DE). For 7 of the 10 subsets, accuracy of responding to the subsets exceeded the level predicted by chance and was uniformly high across all subsets. The 3 exceptions were subset BC, BD and CD. These data, which are shown in Figure 3, are puzzling for two reasons. The first is that subjects completed the required sequence (A → B → C → D → E) correctly on more than 75% of the trials during a criterial session that occurred a day before the subset test and that the accuracy of responding to BC, BD and CD during that was > 0.9. The second puzzle is the difference in the accuracy of responding to subsets BC, BD and CD between pigeons and monkeys.

The reaction times (RTs) of the responses to the first item of each subset provide important clues as to the manner in

which pigeons and monkeys represented the sequence (A → B → C → D → E) during the subset test. For monkeys, RT to the first item of a subset increased as a function of that item's position in the original list. This suggests that monkeys start to scan their representation of the sequence at item A. If they do not find item A they proceed to item B, and so on. The subjects orderly progression through their representation results in the increasing RT displayed in the left panel of Figure 4.

Once the subject has located the first item, they continue scanning their representation of the list until they identify the second item. For example, after responding to item C in pair CE a subject would proceed to item D in its representation. Upon not finding item D, the subject would continue scanning its representation until it came to item E and would then respond to it. As shown in the right panel of Figure 4, the result of this process is a uniform increase in RT

In contrast to monkeys, pigeons show similar RTs regardless of the location of the first item in the original sequence. Similarly, the number of items between the first and second

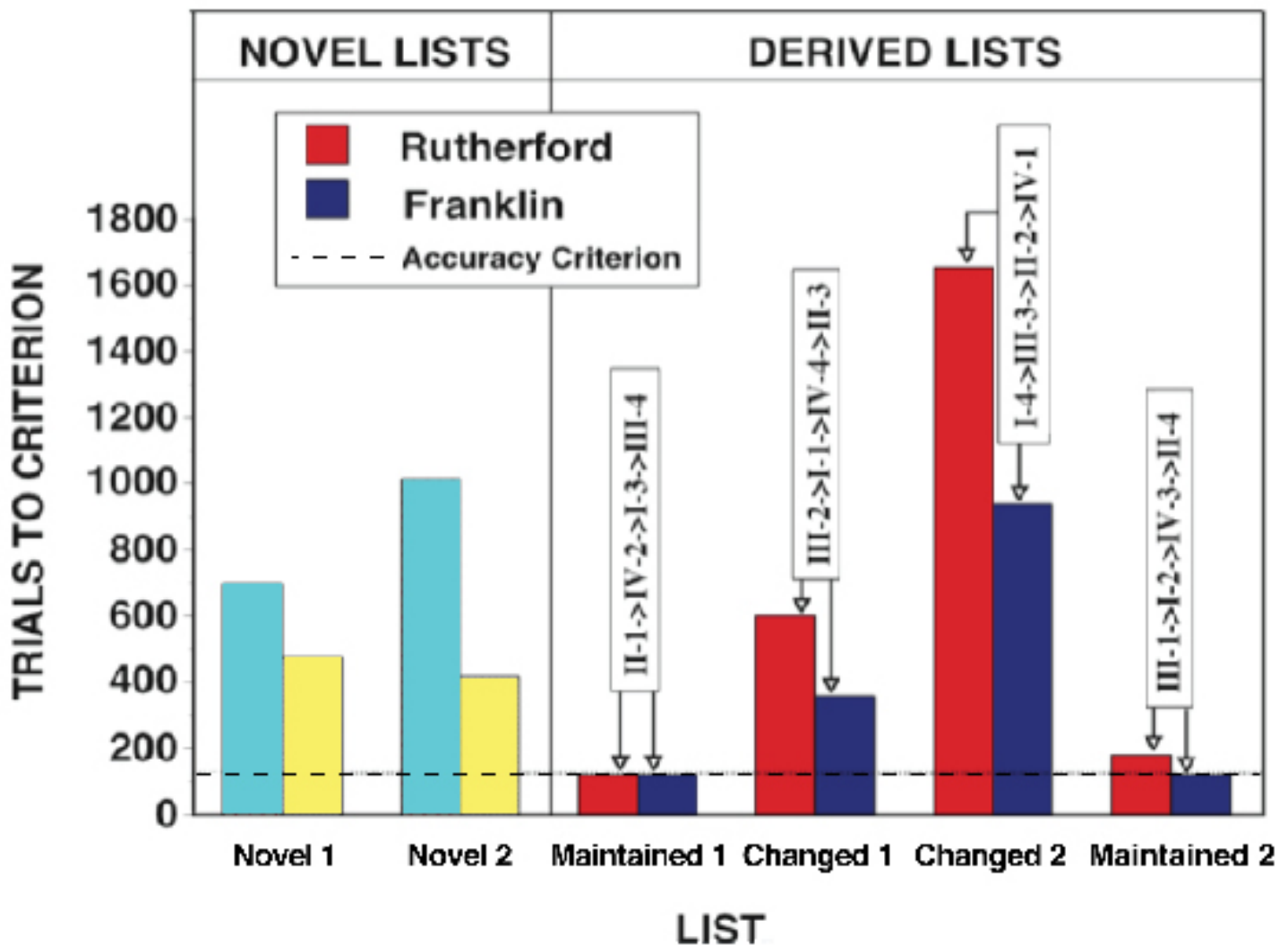


Figure 8. Number of trials needed to satisfy the accuracy criterion on four derived list in the Chen et al. (1997) experiment. The left hand panel shows the number of trials that these subjects needed to learn the last new list on which they were trained. The order in which the derived lists were trained is shown in Figure 7.

item of a pair has no influence on the RT to the second item (not shown). These data suggest that pigeons adopted the following rules when responding on the subset test:

- i. Respond first to item A.
- ii. Respond last to item E.
- iii. Respond to any other item by default.

For pigeons, rules i-iii predict accurate performance on all subsets that contain an end item (AB, AC, AD, AE, BE, CE, and DE). By contrast, they provide no basis for selecting subsets beginning with B, C or D.

Experiments with Monkeys. In the first two experiments on simultaneous chaining in monkeys (D'Amato and Colombo 1988; Swartz, Chen et al. 1991), subjects learned 4 lists by the successive phase method and showed a modest, but steady, increase in the efficiency with which they acquired new lists. This suggested that it might be possible to dispense with the cumbersome successive phase method

entirely, and train lists with all items present from the start of training.

Serial Expertise. In his classic study of learning sets, Harlow showed that it was easier for a subject to induce a general strategy for learning simple cognitive tasks by training it on many versions of a particular problem than to wait for a subject to achieve a high level of accuracy to a particular exemplar of that task (Harlow 1949). This was the approach we followed in an experiment on serial expertise (Terrace, Son et al. 2003), in which monkeys were trained to learn numerous multiple 3-, 4- and 7-item lists on which all items were presented from the start of training (Terrace, Son et al., 2003). We started with 3-item lists because a meaningful list cannot be constructed with fewer than 3 items. A 2-item list does not have to be learned as a sequence because a subject simply has to identify A and then respond to the second item by default. That strategy cannot be used on a 3-item list. After the subject learns A, it has to choose one of the two

Seven-Item Lists

List One



List Two



List Three



List Four



Terrace, Son, Brannon, *Psychological Science* (2003)

Figure 9. Photographs used to compose four 7-item lists on which monkeys were trained in the Terrace, Son, and Brandon (2003) experiment.

remaining items.

With the exception of the first list, each 3- and 4-item list was presented for a maximum of 3 sessions. The idea was to see if a monkey could induce a strategy for identifying the ordinal position of items on new lists. During training on four 7-item lists, subjects were required to meet a criterion of completing at least 65% of the trials correctly on each list in a single session.

The strategy of training multiple exemplars of 3-, 4- and 7-item lists produced an unexpected dividend. Subjects were not only able to learn lists without the benefit of the successive phase method, but they became progressively more efficient at mastering new and longer lists. After training on only seven 3-item lists and eleven 4-item lists, they were able to learn 7-item lists in which all of the items were present from the start of training in progressively fewer trials. As shown in Figure 5, subjects needed an average of 31.5, 17.5, 13 and 12.25 sessions, respectively, to satisfy the accuracy criterion on their 1st, 2nd, 3rd and 4th 7-item lists (ranges:

21-55, 11-25, 11-19, and 7-17, respectively).

Logical Errors. Unless a subject guesses the correct sequence by chance, it must make some errors. These are logical errors that function as hypotheses about the ordinal position of a particular item. If, for example, a subject responds to G at the 2nd position of a 7-item list, that error disproves the hypothesis that G is the 2nd item. If, on some other trial, the subject again responded to G as the 2nd item, that response is considered to be a perseverative error because it provides no new information about the identity of the 2nd item. Instead, it would simply indicate that the subject has forgotten the consequences of the logical error it made on an earlier trial.

The 7-item lists trained in this experiment were only the 19th, 20th, 21st and 22nd lists (of any length) that the subjects had learned. Similar reductions in the amount of training needed to satisfy an accuracy criterion have been observed in experiments in which adult human subjects were trained to memorize successive lists of arbitrarily se-

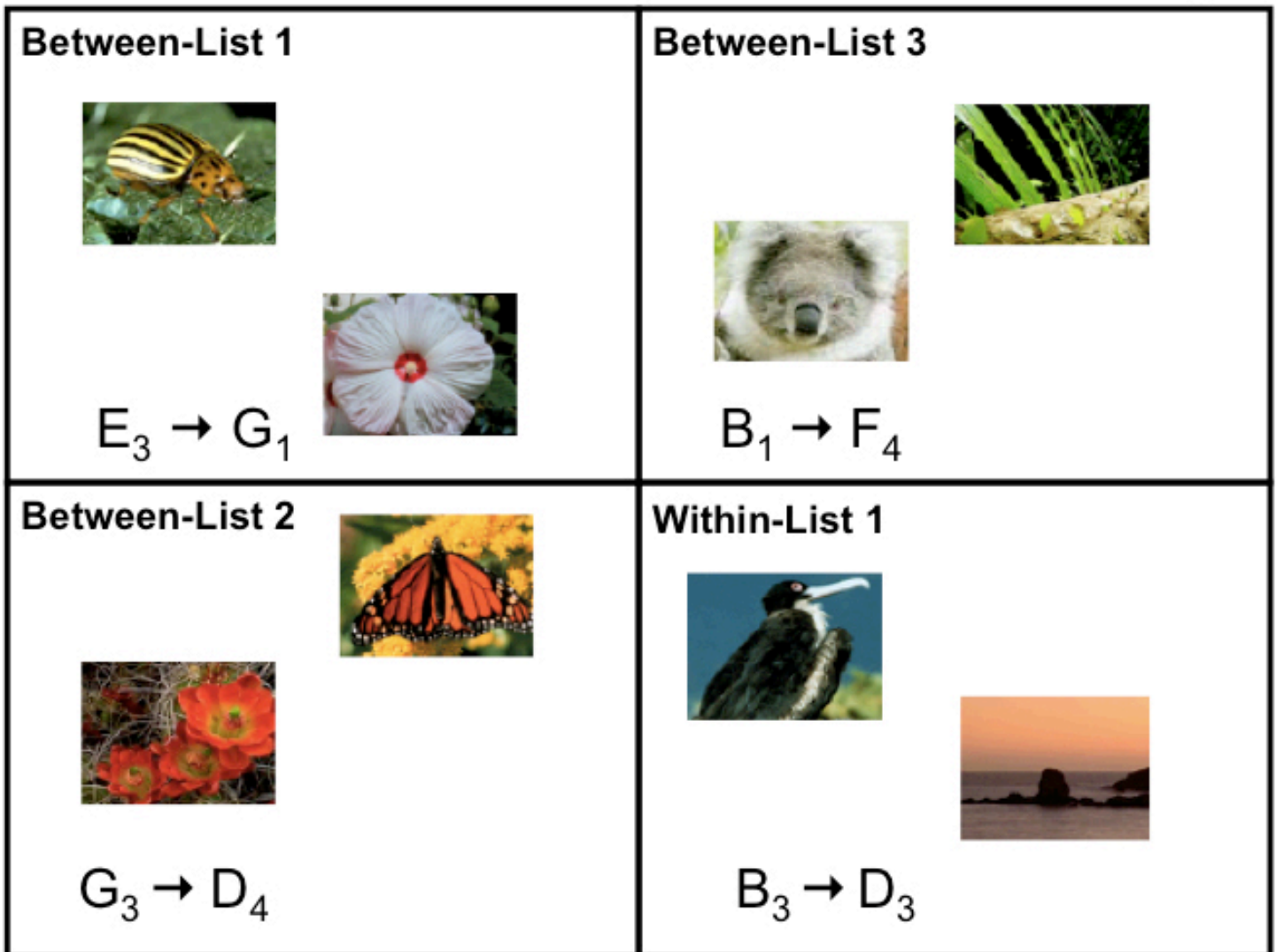


Figure 10. Examples of within- and between-list 2-item subsets used in experiment on serial expertise (Terrace, Son, & Brannon, 2003).

lected words (Keppel, Postman et al., 1968). However, as compared to the monkeys trained in this experiment, human subjects had the benefit of learning thousands of lists prior to their experimental training, not to mention list-learning strategies.

An Ideal List Learner. As shown in Figure 2, it is possible to make 21 different types of forward logical errors at the start of training on a new 7-item list. By definition, each type of logical error can occur only once. An ideal list learner remembers the consequences of each error at a particular position and doesn't repeat that error while learning a new list. If an ideal list learner doesn't guess an item's ordinal position correctly with its first response to that item, its best strategy would be to make logical errors to other items until it encounters the correct item. The average number of logical errors an ideal list learner would make while learning a new n -item list is the product of the maximum number of logical errors that can be made at a given position and the probability of guessing the correct item at that position. For

example, the maximum number of logical errors needed to determine A on a 7-item list is $6 + 5 + 4 + 3 + 2 + 1 = 21$, and the probability of a correct logical guess to A = $1/7$. Thus, the expected number of logical errors needed to determine A is $21 * 1/7 = 3$. Similarly, the maximum number of logical errors needed to determine B is $15 * 1/6 = 2.5$; the maximum number to determine C is $10 * 1/5 = 2$, and so on. The value of the expected number of logical errors at each position decreases linearly, by 0.5 guesses at each position, until it reaches a value of 0 at item F. It is not possible to make logical error at F, the 7th position, because F can be identified by default.

Accuracy on Partially Completed Trials. The functions shown in Figure 5 underestimate subjects' serial knowledge because they are based entirely on correctly completed trials. The conditional probability of responding correctly to each item is a more sensitive measure of serial knowledge because it provides credit for partially correct trials. In contrast to an overall measure of accuracy, which assigns a single

value to the outcome of each trial, conditional probabilities assign an equal weight to each correct response on a given trial.

Figure 6 shows three conditional probability functions that provide additional evidence of subjects' serial expertise. Each function depicts subjects' performance during the first session of training on 3 new lists: on the last 4-item list, and on the first and the last 7-item lists. The probability of responding correctly to all of the items on a given trial is the product of the conditional probabilities of responding correctly to each item. For example, if the conditional probabilities of a correct response to A, B, C, D, E, F and G on the first day of training on the first 7-item list were, respectively, .79, .62, .49, .46, .32, .27 and .29 but the relative frequency of a correctly completed trial was only .003.

The extent to which the number of logical errors a subject approximates that of an ideal list-learner is a measure of the subject's expertise at learning new lists. On a 7-item list, an ideal list learner would need, on average, 3 logical errors to identify A; 2.5 logical errors to identify B; 2 logical errors to identify C, and so on. Given 60 trials/session, and an average of 3 logical errors that would be expected at A, the conditional probability of a correct response by an ideal ordinal position detector at position A would be $57/60 = .95$; given an average of 2.5 logical errors at position B, $54/57 = .96$.

At the start of training on their 4th 7-item list, a comparison of the performance of an ideal list learner with that of the subjects of the Terrace, et al. experiment provides powerful evidence of their serial expertise. In particular, subjects identified items that occupied the positions A and B almost as accurately as an ideal list-learner would. On average, subjects made only 4 errors at position A (accuracy = .92) and 5 errors at position B (accuracy = .88) during the first session of training on that list. Thus, at the start of training on their 4th 7-item list, subjects were able to deduce the position of the first two items with almost perfect efficiency.

Knowledge of Ordinal Position. The availability of list-sophisticated monkeys makes it possible to ask questions about the serial organization of memory that, previously, could only be addressed in experiments on human subjects. That opportunity was exploited in 2 experiments (Chen, Swartz et al., 1997; Terrace, Son et al., 2003). The first experiment is a replication of a classic experiment on knowledge of the ordinal position of list items in human subjects (Ebenholtz 1963) that used Ebbinghaus' method of derived lists (Ebbinghaus 1964). The second is an extension of the expertise experiment described earlier.

Derived Lists. The subjects of this experiment were two male monkeys that had been previously trained to produce

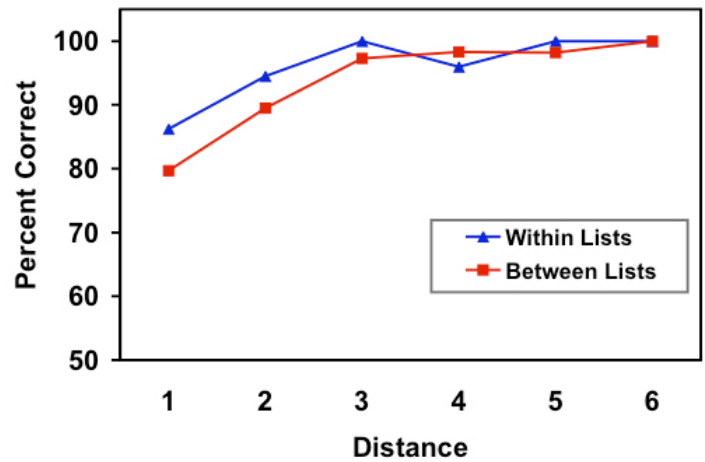


Figure 11. Mean accuracy of responding to between- and within-list subsets as a function of distance between items on original lists in Terrace, Son, and Brannon (2003) experiment.

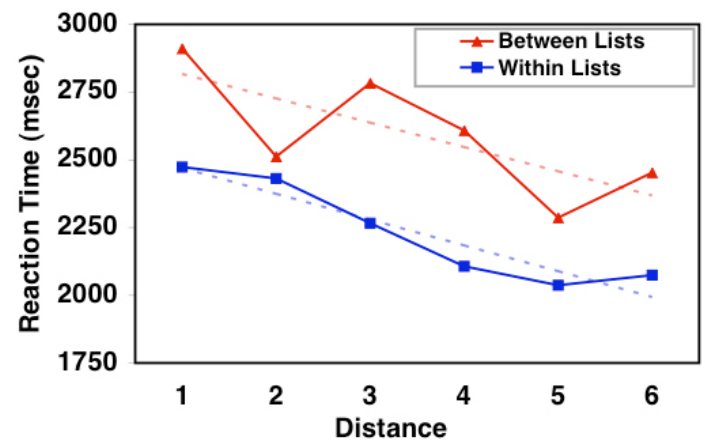


Figure 12. Mean RTs of monkeys responding to between- and within-list subsets as a function of distance between items on original lists (Terrace, Son, and Brannon, 2003).

four 4-item lists at a high level of accuracy (Swartz, Chen et al., 1991). The transfer lists we used were constructed by the method of derived lists (Ebbinghaus, 1964). Subjects were trained to produce four derived lists that were composed of items from the four lists they had mastered previously (one item from each of the 4 original lists). The latter are designated as: $A1 \rightarrow B1 \rightarrow C1 \rightarrow D1$, $A2 \rightarrow B2 \rightarrow C2 \rightarrow D2$, $A3 \rightarrow B3 \rightarrow C3 \rightarrow D3$ and $A4 \rightarrow B4 \rightarrow C4 \rightarrow D4$. As shown in Figure 8, each derived list was constructed with the constraint that only one item could be derived from one of the 4 lists on which they were trained earlier. Thus, each derived list contained one, and only one, item from each of the first 4 lists that these subjects learned. The ordinal position of each item was maintained on two of the derived lists (Maintained Lists 1 & 2: $A2 \rightarrow B4 \rightarrow C1 \rightarrow D3$ & $A3 \rightarrow B1 \rightarrow C4 \rightarrow D2$, respectively). On the other two, (Changed Lists 1 & 2), the original ordinal position of each item was changed ($B3 \rightarrow A1 \rightarrow D4 \rightarrow C2$ & $D1 \rightarrow C2 \rightarrow B3 \rightarrow A4$). The original

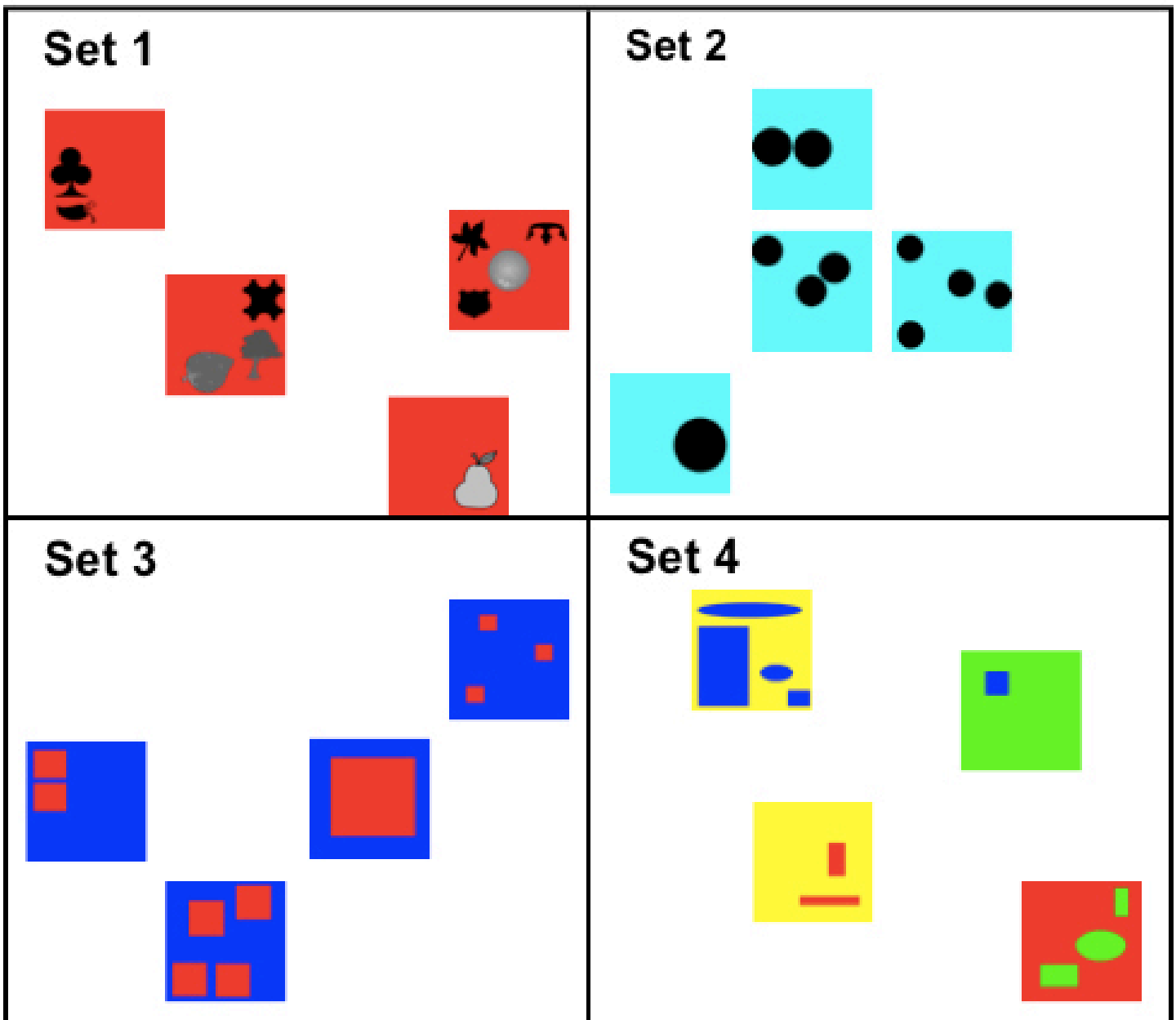


Figure 13. Exemplars of the different types of numerical stimuli on which monkeys were trained in the Brannon & Terrace 1998 experiment. Experimental stimuli varied with respect to the size, shape and color of the geometric stimuli from which they were composed.

and the derived lists are shown in Figure 7.

When a subject satisfied the accuracy criterion for its 4th original list, it was advanced to its 1st derived list. Given that each derived list contained only one item from each of the previously learned lists, all lists should be equally difficult if a monkey's knowledge of the 4 original lists was limited to item-item associations. If, however, a monkey learned an item's ordinal position on the original lists, the 2 maintained lists should be acquired more rapidly than the 2 changed lists. The performance of both subjects confirmed the latter prediction. Transfer was almost perfect for the 2 maintained lists. Both monkeys acquired Maintained List 1

in the minimum number of trials needed to satisfy the accuracy criterion (120). On Maintained List 2, subject 1 satisfied the accuracy criterion in 120 trials, while subject 2 required 180 trials. By contrast, both subjects had difficulty learning Changed List 1 and Changed List 2, in particular Changed List 2, the derived list on which the ordinal positions of the two end items were switched. The relevant data are shown in Figure 8.

Subset Tests. A second method for assessing a monkey's ordinal knowledge used pairs of items that were selected from a simultaneous chain that had been mastered previously. In one experiment, the arbitrary list items were chosen

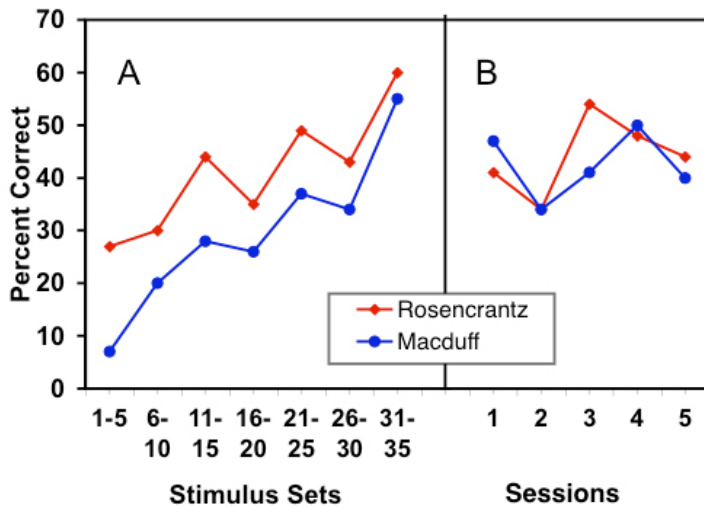


Figure 14. (A) Percentage of correctly completed trials by monkeys during the first session of each of the 35-training stimulus sets (blocks of five sessions). (B) Percentage of correctly completed trials on the 150 test sets, each composed of trial unique stimuli (Brannon & Terrace, 1998).

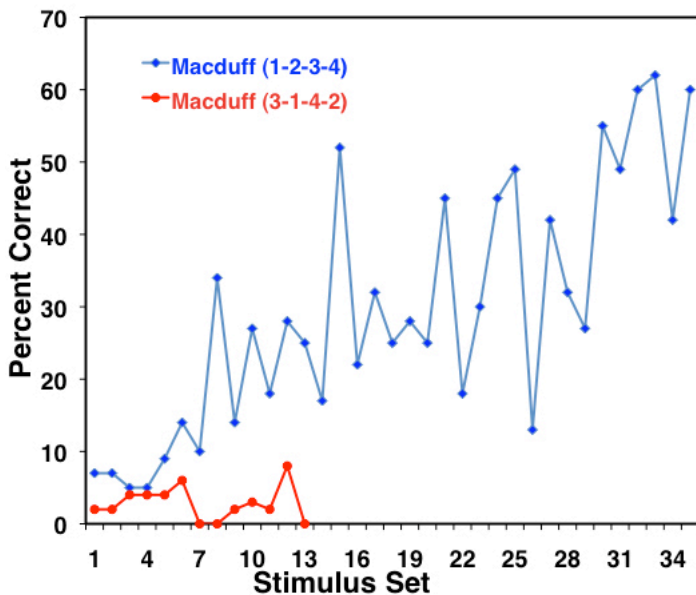


Figure 15. Comparison of one monkey's performance on a non-monotonic (3→1→4→2) and a monotonic (1→2→3→4) simultaneous chain composed of numerical stimuli (Brannon & Terrace, 2002).

from the four 7-item lists that subjects mastered in the expertise experiment (Terrace, Son et al. 2003). These are shown in Figure 9. Subjects were tested on the 28 photographs that were used to compose the four 7-item lists trained in the Terrace et al. (2003) experiment. From these photographs, we selected all possible 2-item subsets, 366 all told. Eighty-four subsets were composed exclusively with items from individual lists, e.g., the subsets A3B3, A3C3...A3G3; B3C3, B3D3...F3G3, from List 3. From the monkey's point of view, the order of the photographs on each of the four lists is just as arbitrary as the order of the letters in any of the four

alphabets. Thus, in the case of the English alphabet, there is no inherent reason why D has to come after C. A child would find either order equally difficult to master. Because a child that knows the alphabet also knows numbers, a child has an important advantage over the uneducated monkeys used in the Terrace, et al (2003) experiment. A child could use numerical symbols as a mnemonic for remembering that letter's position in the alphabet from which it was drawn. There is no reason to think that a monkey could use that strategy. The 252 between-list subsets, were composed of items drawn from different lists, e.g., the subsets A2B4 from Lists 2 & 4, C3F5 from Lists 3 and 5, E1G3 Lists 1 and 3, etc. The subset test was administered during the course of 4 successive sessions.

For the purpose of analysis, subsets were assigned to one of 6 categories on the basis of the distance between their ordinal positions on the original lists (e.g., pairs of items separated by a distance of 1: the subsets A1B1, B2C2...A1B2; B2C3, C3D4...F3G4; a distance of 2: A1C1, B2D2...A1C2; B2D3, C3E4...E3G4; a distance of 6: A1G1, A2G1...A2G1, A2G2, A2G3...A4G4). Within- and between-list pairs were interspersed randomly throughout the subset test. Subjects were rewarded if they responded in the order specified by their ordinal positions on the original list. For example, subjects were rewarded for responding to items C and F, in that order, whether or not they came from the same or different lists (e.g., C3F3 or C2F6). Some examples of within- and between-list subsets are shown in Figure 10.

Predictions of the outcome of this experiment that are based on spatial representations of each list differ dramatically from those that are predicted by associative theory. Theories of serial learning based on cognitive maps or spatial representations of lists would predict that accuracy should increase as the distance between the items increases and that RTs should decrease because larger distances are more discriminable than smaller distances (Holyoak and Patterson, 1981). Associative theory would predict that accuracy of responding to subset pairs would decrease as the distance between the items increased and that RTs would increase because the number of associative steps needed to determine the ordinal position of each item increases with distance.

Predictions based on a spatial representation of a simultaneous chain were confirmed. The relevant data are shown in Figure 11 and 12. Remarkably, accuracy was close to 100% during the first session on which a particular subset was displayed. None of the differences in the accuracy of subjects' responses to within- and between-group subsets were statistically different.

Numerical Sequences. Suppose that you were asked to train a monkey to execute the sequences 1-2-3-4-5-6-7, a set

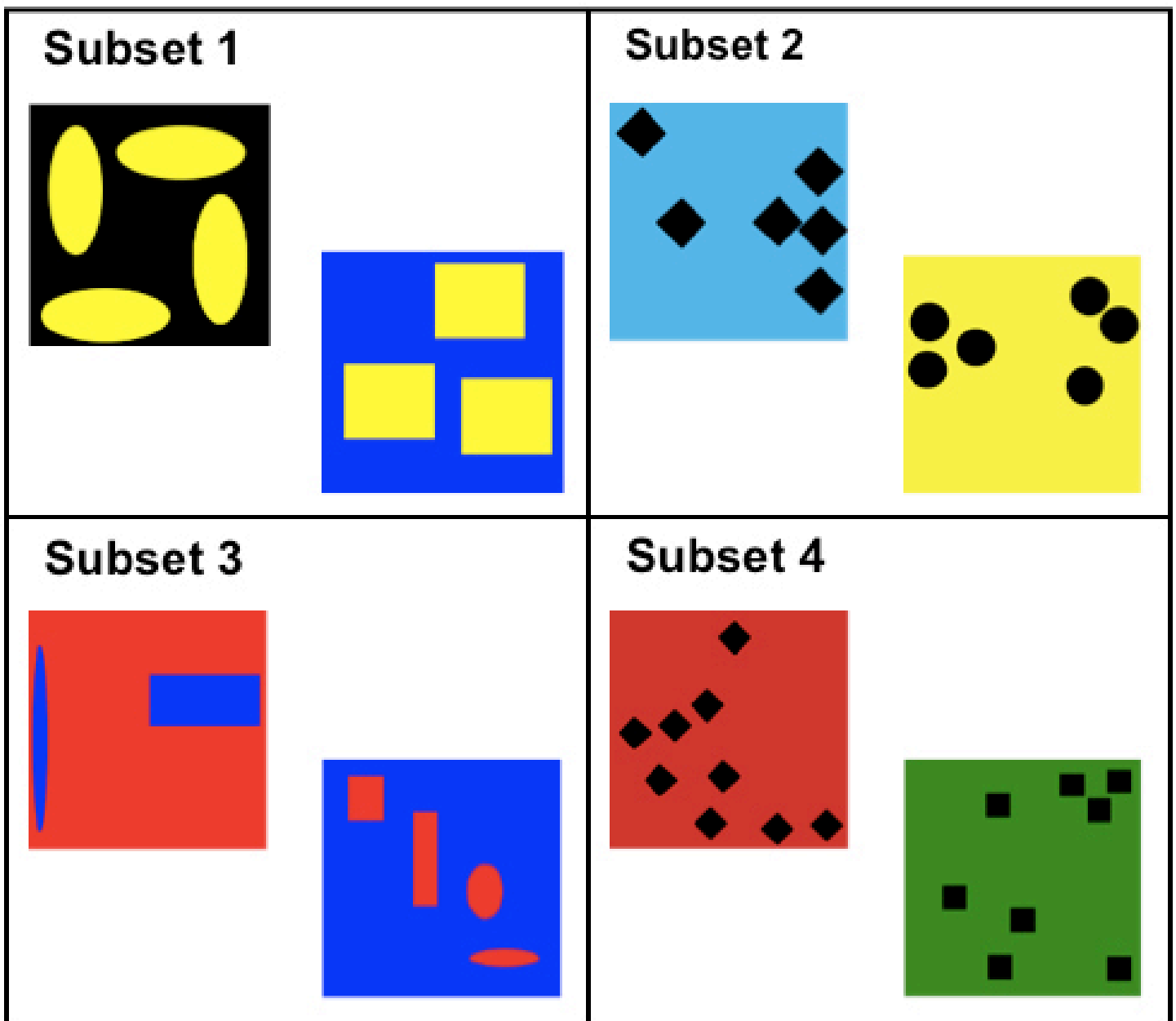


Figure 16. Exemplars of subsets used as novel numerical stimuli (Brannon & Terrace, 1998).

of photographs or A-B-C-D-E-F-G. There should be no difference in the difficulty of these sequences, unless the monkey had learned that the symbols 1-2-3-4-5-6-7 represent increasing degrees of magnitude where $7 > 6$, $6 > 5$, $5 > 4$, etc. As we have seen in Figure 1B, the numerical properties of the Arabic numerals on a number pad are irrelevant when it comes to dialing a phone number or entering a PIN. Learning to produce the sequence 1-2-3-4-5-6-7 by rote implies nothing about counting because learning to count implies the ability to discriminate sets of objects on the basis of their numerosity and knowledge of the ordinal and cardinal properties of numbers (Gelman and Gallistel, 1978).

Nominal, ordinal and cardinal refer to three different properties of numbers. Consider the following 3 usages of

the Arabic Numeral 24. In connection with the TV program that bears its name or a baseball player that wears 24 on his back, 24's only function is to distinguish that TV program from others or that baseball player from other players, hence the term "nominal". A meaningful way to distinguish a baseball player is his rank, i.e., his ordinal position, with respect to some skill, also an Arabic numeral. In this instance we can distinguish 24 from another baseball player by comparing their batting averages, the number of stolen bases, etc. Any ordered series of symbols would work just as well, e.g., the alphabet, Roman numerals, etc. Since the distance between people or items that are ranked is unknown, adding or subtracting ranks has little meaning if a person's rank is all you know about him. How, for example would you interpret the ranks of the top two students in the senior class if their GPAs

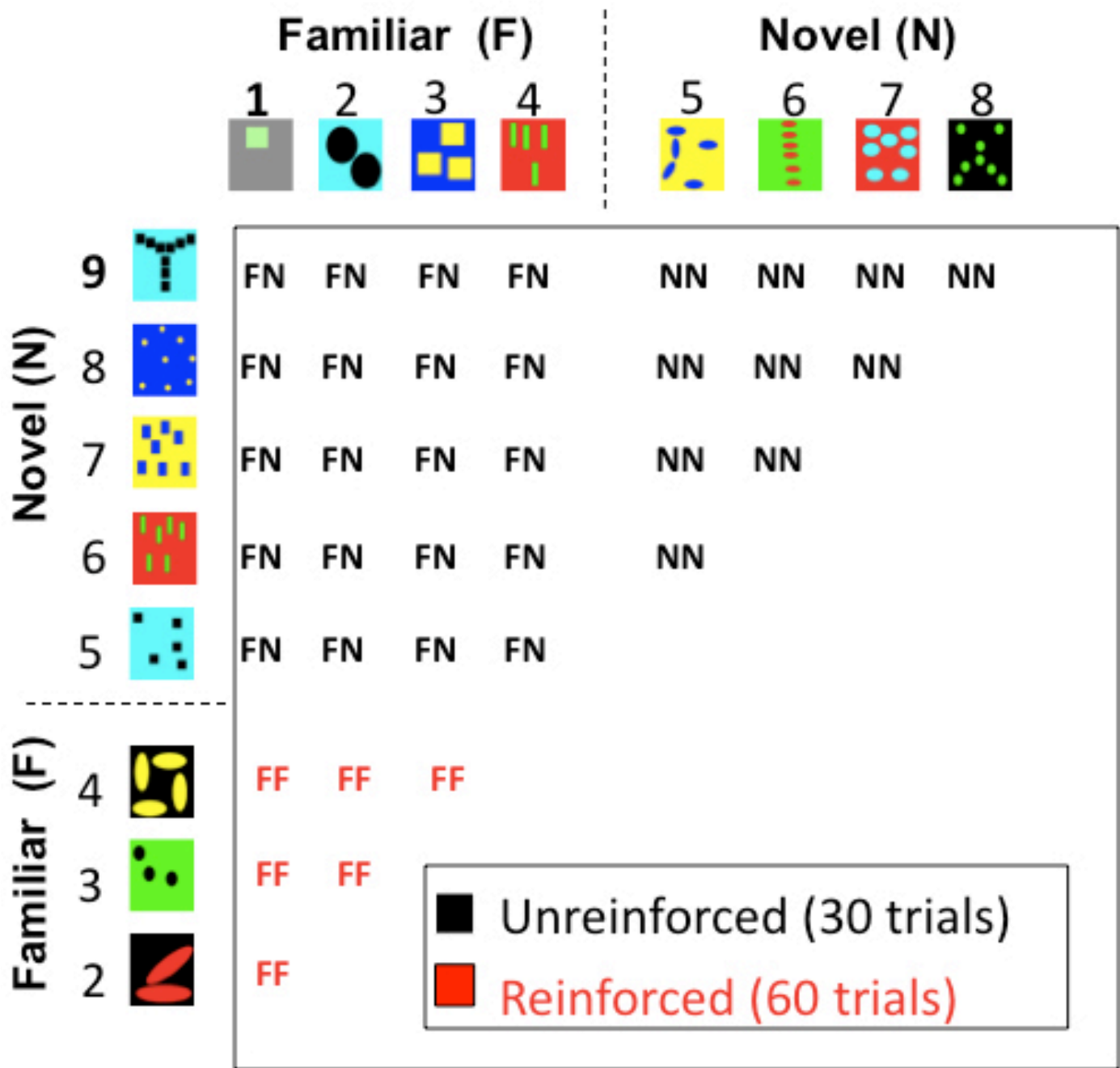


Figure 17. The 36 pairs of the numerosities used in the pair wise test (1-9) (Brannon & Terrace, 1998). These pairs were segregated into three categories that were defined with respect to the subjects' prior experience with the constituent numerosities: familiar-familiar (FF), familiar-novel (FN), or novel-novel (NN).

were, respectively, 3.86 and 3.85? A different interpretation would apply if their GPAs were, respectively, 3.92 and 1.63. Thus, a college admission officer that is evaluating different students that had the 2nd highest GPA in their graduating class needs more information before she can make a valid decision about admitting one of those students.

Cardinal knowledge implies two other properties of numbers. To say that you have cardinal knowledge of the Arabic number 3 is to say that you could use that symbol appropriately whenever you see a set of 3 (and not 4 or 5, etc.)

objects. Understanding that the Arabic number 3 is exactly one unit bigger than Arabic numeral 2 or that it is one unit smaller than Arabic numeral 4 is the other property of cardinal knowledge. There is much evidence that a non-human primate can learn the ordinal properties of a number, but evidence about cardinal knowledge is controversial.

Because animals do not normally count, claims about their knowledge of the ordinal and cardinal properties of numerical symbols have been met with considerable skepticism. Accordingly, experiments on the numerical ability

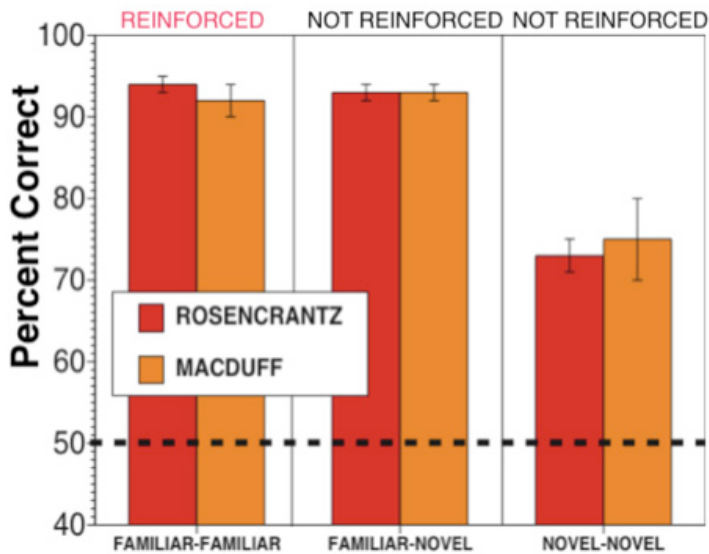


Figure 18. Both subjects' accuracy exceeded the chance-level accuracy on Familiar-Familiar, Familiar-Familiar and Novel-Novel pairs (Brannon & Terrace, 1998).

of an animal have had to bear the burden of proof when it was claimed that the animal in question lacked any of the components of numerical ability that are natural in children (Davis and Perusse 1988; Roberts, 1998). That burden notwithstanding, investigators of animal behavior have shown that many species do have a limited degree of numerical ability. Those observations have led some psychologists to hypothesize that human mathematical ability evolved from numerical abilities that can be observed in animals (Gallistel and Gelman, 1992; Dehaene and Cohen, 1997; Gallistel and Gelman, 2000).

Brannon and Terrace (Brannon and Terrace 1998; Brannon and Terrace 2000) have collected considerable evidence of the ordinal abilities of rhesus monkeys that support the continuity hypothesis. A basic assumption of those experiments was that a monkey must be trained to discriminate the numerosity of different sets of objects before it can learn the ordinal or the cardinal properties of numbers. To demonstrate ordinal ability, monkeys were trained to acquire simultaneous chains composed of numerically defined stimuli. Those stimuli, which were composed of geometric forms that varied in color, shape, number, and size, could be perceived as arbitrary shapes or as exemplars of a particular numerosity. Some examples of the numerical lists used in these experiments are shown in Figure 13.

The first goal of the Brannon and Terrace (1998) experiment was to show that monkeys could learn a list composed of ascending numerical stimuli more rapidly than a list composed of arbitrarily selected photographs. Accordingly, they trained two monkeys on 35 four-item lists composed of exemplars of the numerosities 1, 2, 3 and 4 in an ascend-

ing order. Brannon & Terrace (2002) have also successfully trained monkeys to learn a numerical list on which the rule was to respond in a descending order (4-3-2-1) but showed that there were basic differences in performance between subjects that learned ascending and descending lists. These will be discussed later.

Exemplars of the numerosities 1, 2, 3, and 4 were selected from a large library of stimuli in which area, shape, size and color were varied systematically to insure that they could not serve as cues for numerosities. Each of the numerical lists on which the subjects were trained was constructed from novel exemplars of the numerosities, 1, 2, 3 or 4. Following Harlow's procedure for establishing learning sets, each stimulus set was presented, until the subject responded correctly during two successive sessions on at least 30% of the trials during a particular session or, until it completed 3 sessions on each list. After training on 35 stimulus sets, subjects were tested on novel lists to see if they abstracted a numerical rule

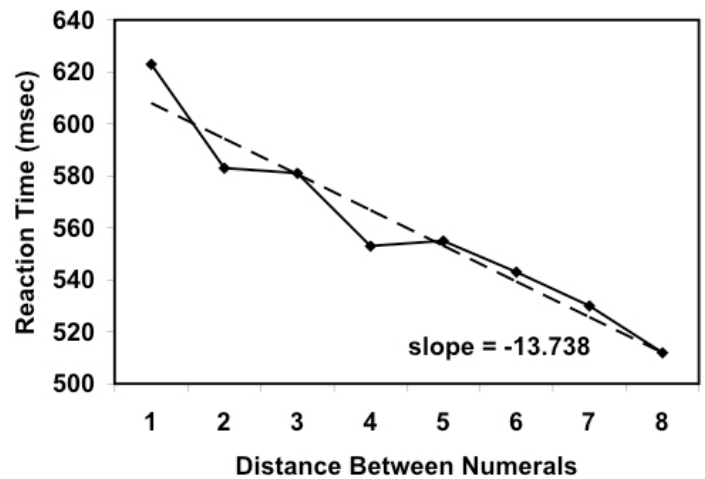


Figure 19. RTs of human subjects asked to select the larger of two simultaneously presented Arabic numerals (Moyer & Landauer, 1967).

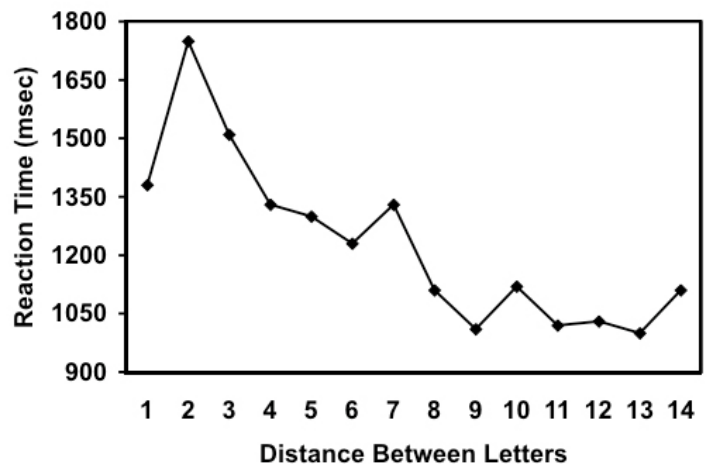


Figure 20. RTs of human subjects asked to select which of two simultaneously presented letters come first in the English alphabet (Hamilton & Sanford, 1978).

during their initial training.

Memory of Particular Stimuli or of Numerical Categories? To rule out the unlikely possibility that the monkeys memorized each of the 35 lists on which they were trained, they were tested on their ability to respond in the correct order to 150 new lists. Each of those lists was composed of trial unique stimuli and was used for only one trial. Each list was presented for only one trial. Subjects continued to respond at approximately the same level of accuracy (40%) to the trial unique lists as they did at the end of training on the 35 original lists. The slight decrement in accuracy of responding during training on lists composed of novel exemplars of the numerosities 1-4 was not statistically significant (cf. Figure 14). In this connection it should also be noted that probability of responding accurately on a novel 4-item is $< .05$. Since subjects had no opportunity to memorize any of the novel lists, the absence of a performance decrement provides clear evidence that subjects used the numerosity of each stimulus (as opposed to particular physical features) to determine the order in which to respond on the novel lists. The relevant data are shown in Figure 14.

Less clear is how a monkey represents the numerosities 1-4. One possibility is that the monkeys assigned each numerosity to one of four distinct nominal categories. Under this scenario, subjects would have learned an arbitrary ordering of the four categories, just as if they learned to respond to exemplars of, say, birds, flowers, trees, and rocks. If that were the case, it shouldn't matter if the simultaneous chain on which they were trained was composed of numerosities such as, 3-1-4-2, as opposed to 1-2-3-4. Contrary to that hypothesis, the one subject that was trained on a 3-1-4-2 sequence showed no signs of improvement after it was trained on 13 stimulus sets. However, as can be seen in Figure 15, that subject's performance improved immediately once it was required to respond to list items in an ascending order. The relative ease of learning a monotonic rule provides strong evidence that monkeys perceive the ordinal relations between the numerosities on which they were trained.

Given that evidence, the same subjects were used to evaluate a monkey's ability to perceive ordinal relations between novel numerosities. Specifically, the question Brannon & Terrace asked was, could a monkey apply the ascending rule that they learned by virtue of their training on the numerosities 1-4, to the novel numerosities 5-9? To answer that question, subjects were tested on exemplars of all possible pairs of the numerosities 1-9. Sample subsets are shown in Figure 16. As shown in Figure 17, each pair was classified as familiar-familiar, familiar-novel and novel-novel. Reinforcement for responding to pairs in an ascending order was only available on trials on which one of the 6 familiar-familiar

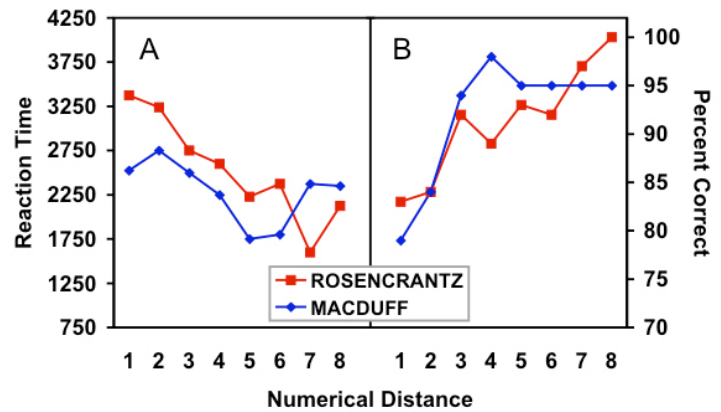


Figure 21. (A) Mean RTs; (B) mean accuracy of responding to 2-item numerical subsets as a function of numerical distance between the numbers of each pair in the Brannon and Terrace (2000) experiment.

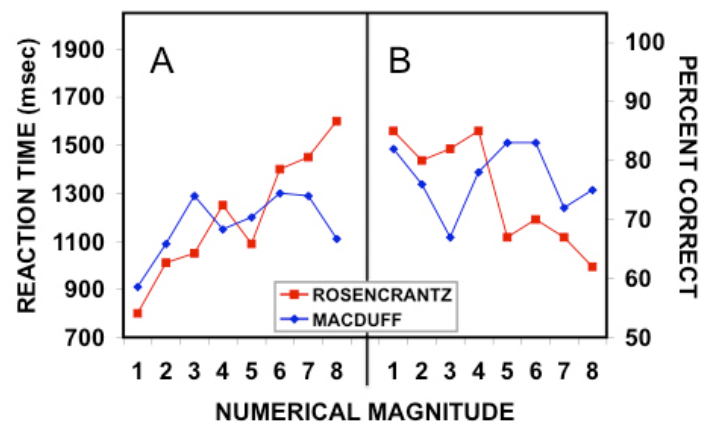


Figure 22. (A) Mean RTs and (B) mean accuracy of responding to 2-item numerical subsets as a function of the numerical magnitude in the Brannon and Terrace (2000) experiment.

pairs were presented (1-2, 1-3, 1-4, 2-3, 2-4, 3-4). Accordingly, there was no basis for learning the ordinal relationship of a novel numerosity with any other numerosity, novel or familiar.

Subjects responded correctly on approximately 75% of the trials on which novel-novel pairs were presented. The relevant first trial data are shown in Figure 18. Both subjects were just as accurate on pairs for which the larger numerosity covered a small area than for pairs for which the larger numerosity covered a larger area. Similar controls for color, shape and size ruled out non-numerical features of the test stimuli as an explanation of subjects' accuracy. Since neither subject had any previous training with the numerosities 5-9, their ability to respond to those numerosities in an ascending order provides clear evidence that a monkey can extrapolate an ascending rule to novel numerosities.

Symbolic Distance and Magnitude Effects. Distance and Magnitude Effects in are typically obtained in experiments

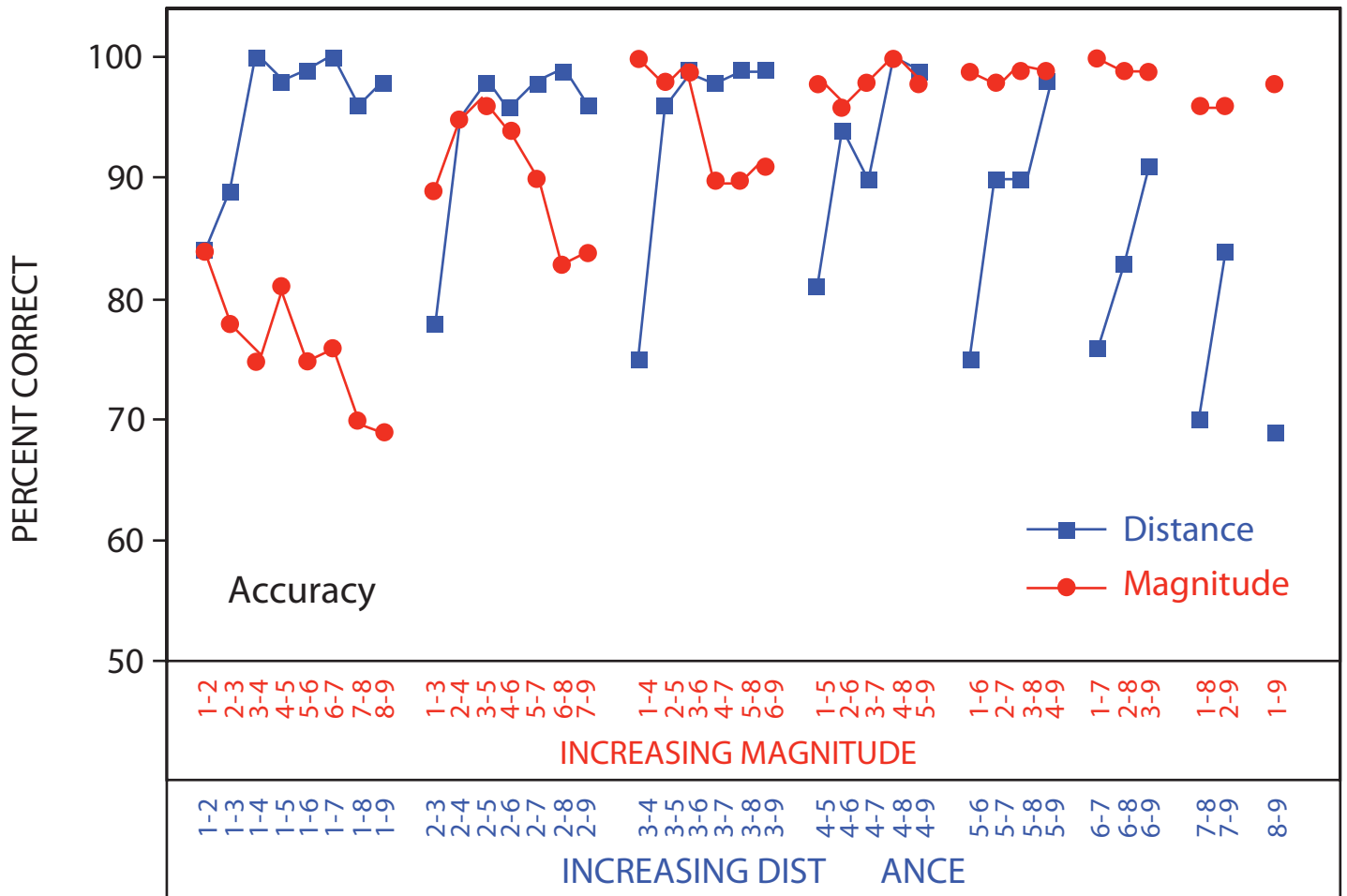


Figure 23. Mean accuracy of responding of monkeys' responses to each of the 36 numerical pairs presented when distance (gray functions) and magnitude (black functions) were held constant (Brannon and Terrace, 1998).

in which human subjects are asked to judge the relative magnitude of two stimuli from some physical continuum, e.g., line length, light intensity, sound energy, etc. (Herrnstein and Boring 1965). Accuracy increases and reaction time (RT) decreases as the ordinal distance between two items increases. The same effects have been observed when human subjects are asked to judge of the relative magnitude of stimuli from psychological continua, for example, the relative magnitude of pairs of Arabic numerals (Moyer and Landauer 1967). As shown in Figure 21, reaction time (RT) decreases as the numerical distance between the two Arabic numerals increases. Moyer & Landauer referred to the inverse relationship between distance and RT as the symbolic distance effect (SDE) because, in this instance, distance is defined psychologically, as opposed to physically. As shown in Figure 20, SDEs have also been obtained in judgments of the ordinal distance along arbitrary continua, e.g., the alphabet (Hamilton and Sanford 1978). Thus, when subjects are asked, which comes first, P or Z, their accuracy is higher and their RTs are shorter than when the same question is posed about P and Q.

In addition to the distance between two items, judgments

of the relative magnitude of two stimuli are also affected by the magnitude of the smaller item. Thus, when subjects are asked, which number is larger, 7 or 8, RT is longer than when the same question is asked about the numbers 3 and 4. Similarly, in the case of alphabetical stimuli, RTs to the pair J & K are longer than RTs to the pair C & D. At a constant distance between the smaller and the larger item, RTs of judgments of relative magnitude increase with the magnitude of the smaller item. That relationship, which is identical to Weber's law in the case of natural continua, has been referred to as a symbolic magnitude effect.

Symbolic Distance and Magnitude Effects Obtained from Rhesus Macaques: Numerical Stimuli. The ability of monkeys to learn arbitrary and numerical sequences provided an unprecedented opportunity to assess distance and magnitude effects in a non-human primate. To assess those effects in monkeys that were trained to order arbitrary and numerical stimuli, we analyzed the accuracy and the RT's of subjects' responses on 2-item subset tests. We will first consider the results of subset tests using numerical stimuli.

In one experiment (Brannon & Terrace, 1998), subjects

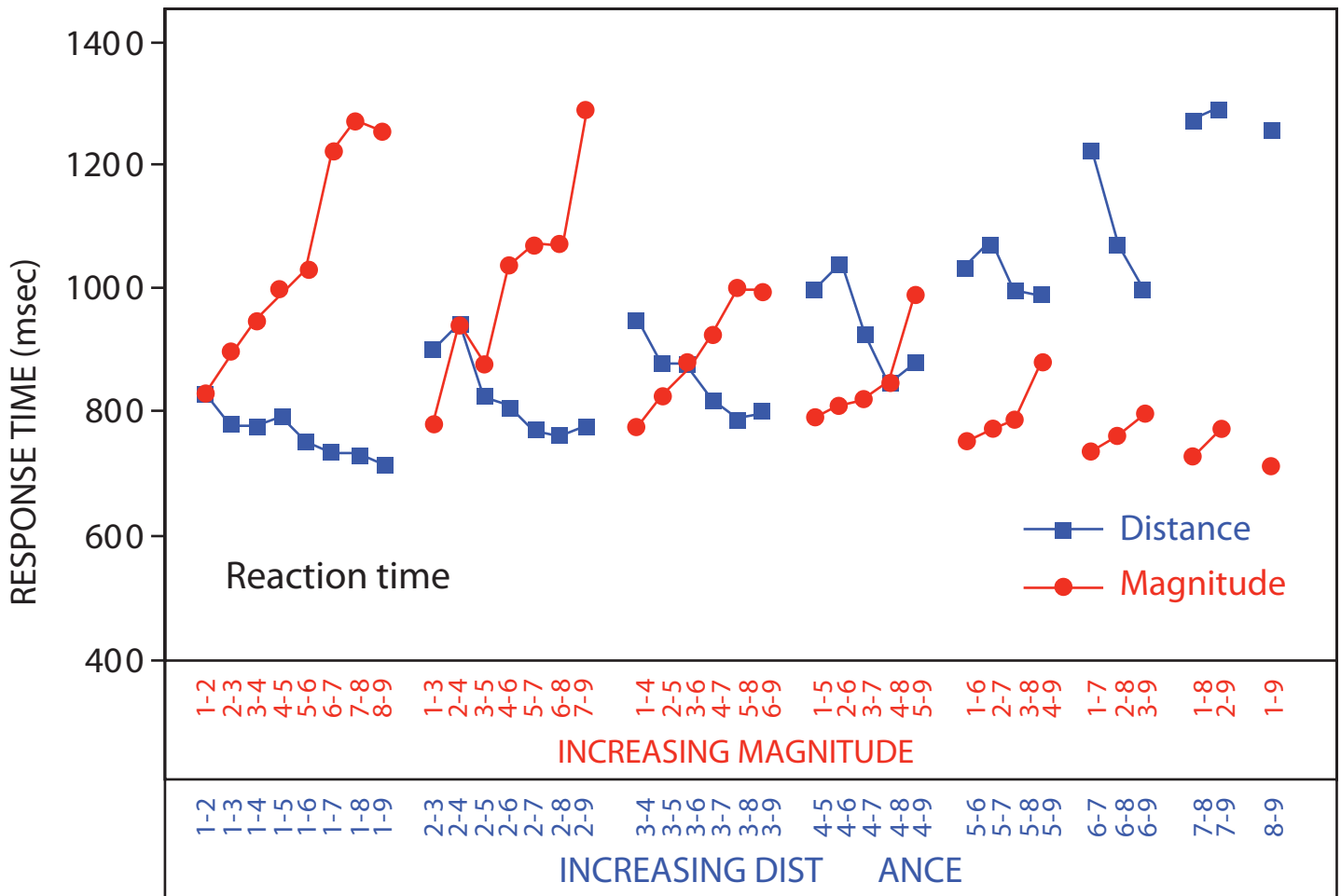


Figure 24. Mean reaction times of monkeys' responses to each of the 36 numerical pairs when distance (gray functions) and magnitude (black functions) were held constant (Brannon and Terrace, 1998).

were tested on the 36 pairs of stimuli that can be drawn from stimuli whose values ranged from 1-9. As can be seen in Figures 21 and 22, there was clear evidence of both symbolic distance and magnitude effects. Accuracy increased and RT decreased as the numerical distance between the members of a pair increased. Analogously, accuracy decreased and RT increased as the magnitude of the first item of each pair increased.

The value of each point shown in Figure 22 is the average of all pairs separated by a distance of one (1-2, 2-3, 3-4, 4-5, 5-6, 6-7, 7-8, & 8-9) a distance of 2 (1-3, 2-4, 3-5, 4-6, 5-7, 6-8, 7-9), a distance of 3 (1-4, 2-5, 3-6, 4-7, 5-8, 6-9), etc. Figure 23 shows the average of all pairs for which the first item had a magnitude of 1 (1-2, 1-3, 1-4...), a magnitude of 2 (2-3, 2-4...), etc. Those averages do not, however, provide information about the uniformity of the effect of distance across all pairs. It is, for example, conceivable that distance might only influence pairs in which the magnitude of the first item exceeded 3.

To determine the uniformity of the effects shown in Figures 21 and 22, those data were analyzed with distance (Fig-

ure 23) and magnitude (Figure 24) held constant. In general, distance and magnitude exerted uniform effects on accuracy and RT. The one exception was the effect of magnitude on accuracy. Were it not for a ceiling effect of 100% accuracy, it is likely that the influence of magnitude on accuracy would have been more pronounced. Figure 23 also shows that the lowest accuracy for any pair of numerosities was significantly greater than the chance level of accuracy. The effect of distance and magnitude on RT was weaker. As can be seen in Figure 24, that effect was most striking for distances 1 and 2.

Arbitrary Stimuli. Data from the subset test using arbitrary list items were analyzed the same way as the data obtained from the subset test with numerical pairs. The average results of that analysis are shown in Figure 25. In each instance, the functions we obtained using arbitrary stimuli bore a strong resemblance to those obtained from subset tests based on numerical stimuli. As was the case with numerical pairs, the functions based on accuracy were compressed because of a ceiling effect. Consider, for example, the upper set of functions in Figure 25 (open squares and triangles). At a distance of 1, the mean level of accuracy was 83%. For distances greater than 2, the average level of accuracy was 99% for

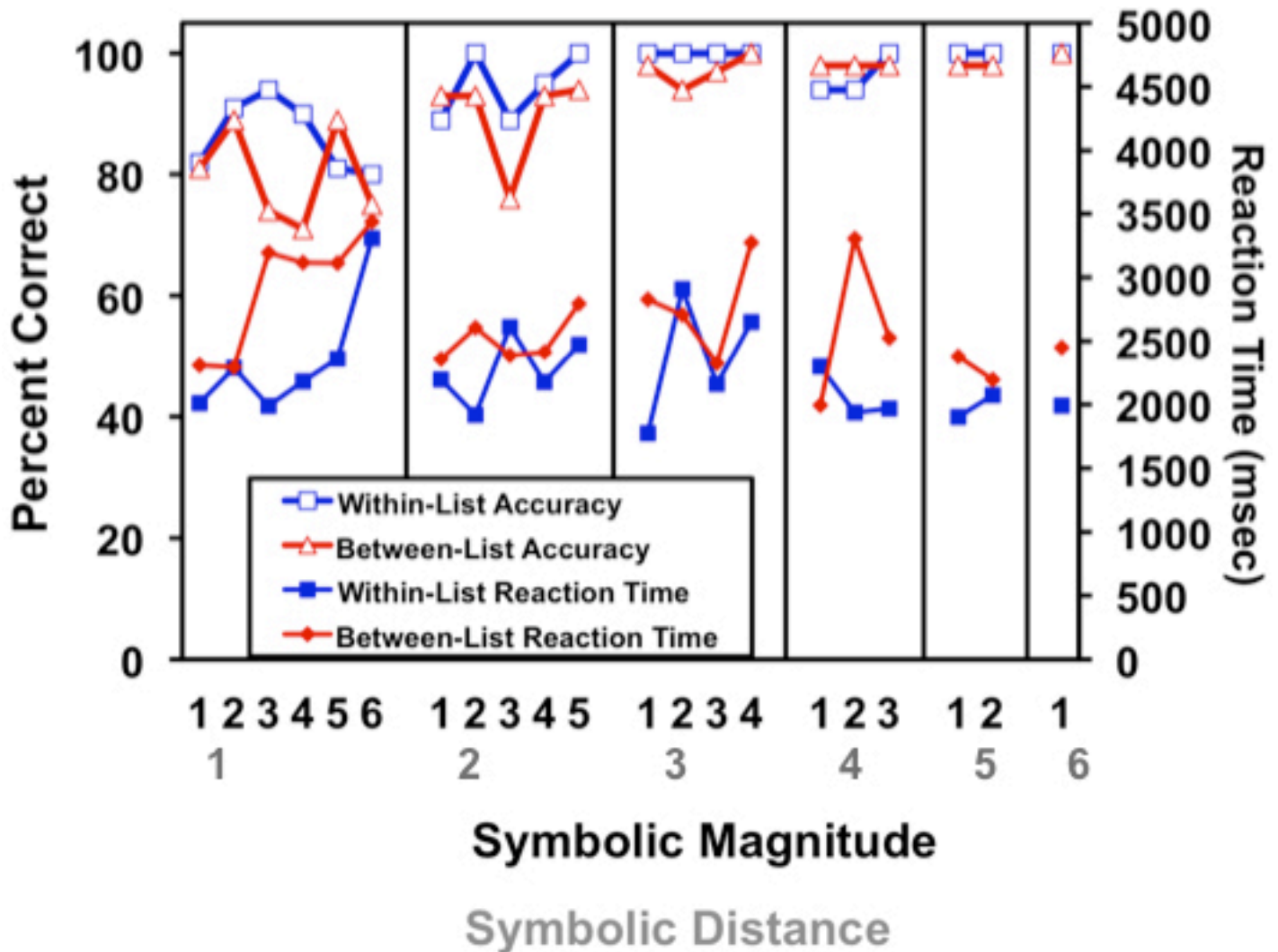


Figure 25. Accuracy and RTs on subset tests. The blue functions show accuracy (top function) and RT (bottom function) on testing within lists; the red functions show accuracy (top function) and reaction time (bottom function) on testing between lists. The entries on the abscissa are generic in that they refer to the types of within- and between-list subsets that are represented at each position. For example, AB refers to A1B1, A2B2, A3B3, and A4B4 in the case of within-list subsets and to A1B2, A1B3, . . . A4B1 subsets in the case of between-list subsets (Terrace, Son, and Brannon, 2003).

all subsets.

The RT data shown in the lower portion of Figure 25 are based on RTs values as function of magnitude when distance was held constant (Figures 23 & 24). They are strikingly similar to analogous RT functions that were obtained with numerical stimuli. For example, at distance 1, RTs increased approximately 470 msec for each increment in ordinal position. For within-group subsets, the median RT to the first item on AB trials was 2,180 msec. On FG trials, it was 3,374 msec. Thus, each increase in the magnitude of the first item of arbitrary pairs increased RT by almost 200 msec. Analogous reaction time functions have been obtained from human subjects in experiments on the discriminability of adjacent pairs of letters of the alphabet when the alphabetical position of the first item is varied from trial to trial (Hamilton and Sanford, 1978; Lovelace and Snodgrass, 1971).

Symbolic Distance and Magnitude Effects in Humans. Because human subjects were tested on such overlearned sequences as the alphabet and a series of Arabic numbers, it was possible that the similarities that were observed were artifacts of the stimuli on which each species was trained. To control for that possibility, experiments were performed on college students that used stimuli that were identical or comparable to those we used with monkeys. One experiment used the same numerical stimuli that were used with monkeys on the same pair-wise comparison task. Subjects were given verbal instructions to respond to the stimulus containing the fewest elements as quickly as possible. As shown in Figure 26, the slopes of the numerical distance and magnitude effects obtained from college students were indistinguishable from those obtained from monkeys (Brannon and Terrace, 2002; cf. Figure 22). These data provide compelling evidence to support the hypothesis that animals

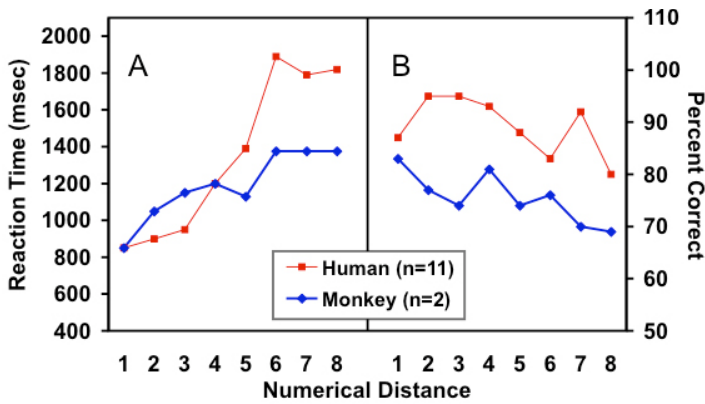


Figure 26. Comparison of performance on 2-item numerical subsets by adult human subjects and monkeys. (A) Mean RT of the first response as a function of distance; (B) Mean accuracy as a function of distance (Brannon & Terrace, 2002).

and humans use the same numerical comparison process.

In another experiment, college students were trained on arbitrary lists composed of complex black and white geometrical shapes of the type shown in Figures 27 (Attneave 1955). Geometrical stimuli, rather than photographs, were used because pilot data showed that human subjects were likely to encode photographs verbally. After subjects were familiarized with the task on a 3-item simultaneous chain, they were instructed to apply the same logic they used on 3-item lists to four 8-item lists. After mastering each list,

subjects were given a 2-item subset test. As can be seen in Figure 28, subjects' RTs resembled those obtained from monkeys. RTs decreased with increasing distance between the subset items and increased as the position of the first item increased.

Similarities in Symbolic Distance and Magnitude Functions in Humans and Monkeys. The similarities in the distance and magnitude effects we have seen thus far are based on the slopes of the functions relating RT to distance and magnitude. For both human and non-human primates, however, the absolute value of RTs was consistently faster for numerical stimuli than for arbitrary stimuli. Differences in the absolute value of RTs to arbitrary and numerical stimuli are evident both within and between species, and indeed, within individual subjects. Consider, for example, the symbolic magnitude functions for numerical and arbitrary stimuli obtained from two monkeys that are shown in Figure 29. RT increased from 800 to 1,300 msec as the magnitude of the first item of a numerical pair increased from 1 to 9. On the subset test that used photographic stimuli, RT increased from 2,200 to 3,425 msec as the position of the first item was advanced within the 7-item list on which the subject were trained.

Representations of Arbitrary and Numerical Lists. Why

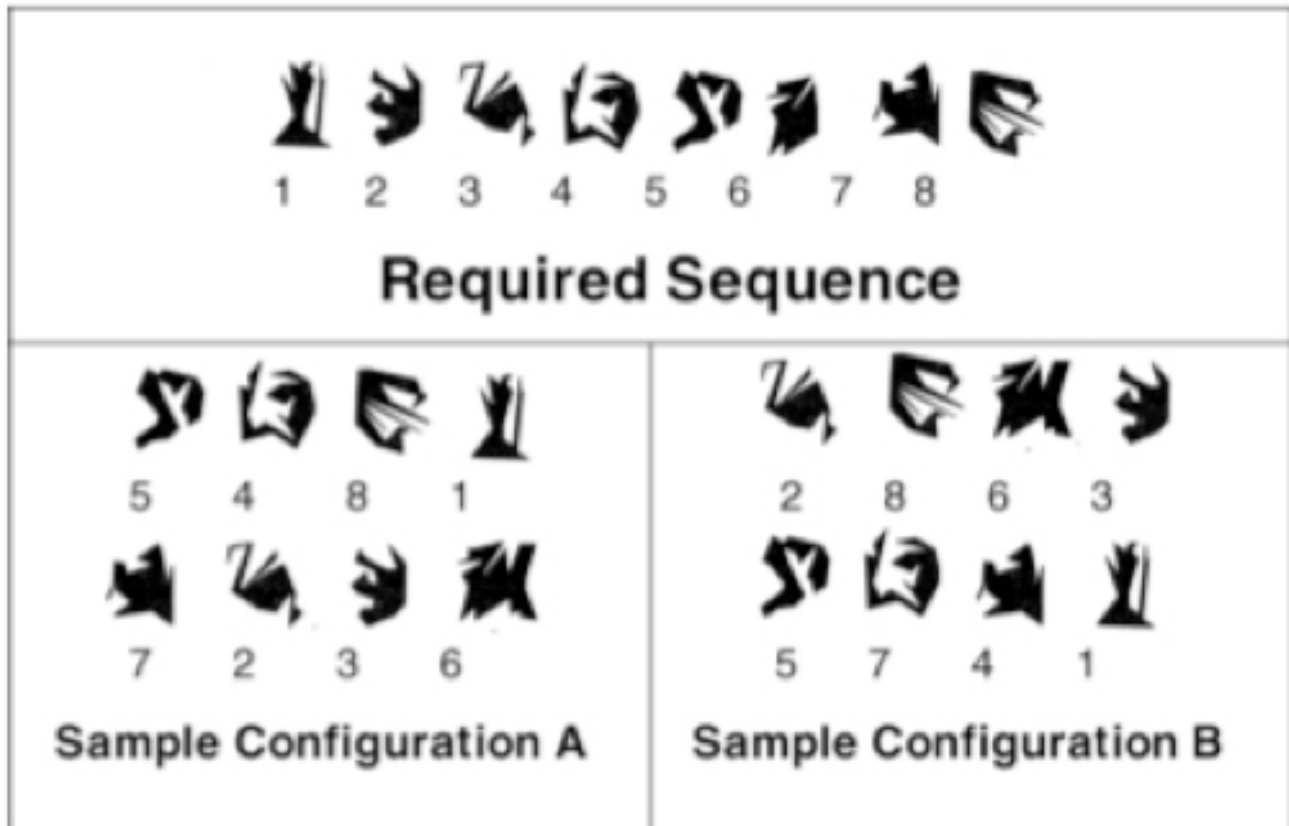


Figure 27. Stimuli used to construct 8-item lists for human subjects trained by the simultaneous chaining paradigm in the experiment by Terrace (2003).

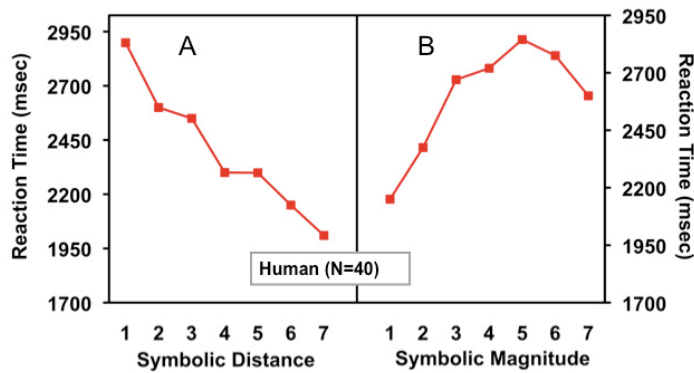


Figure 28. Mean RT of responding of human subjects as a function of the distance (A) and the position (B) of first item on 2-item subsets derived from four 8-item lists (Terrace, 2001).

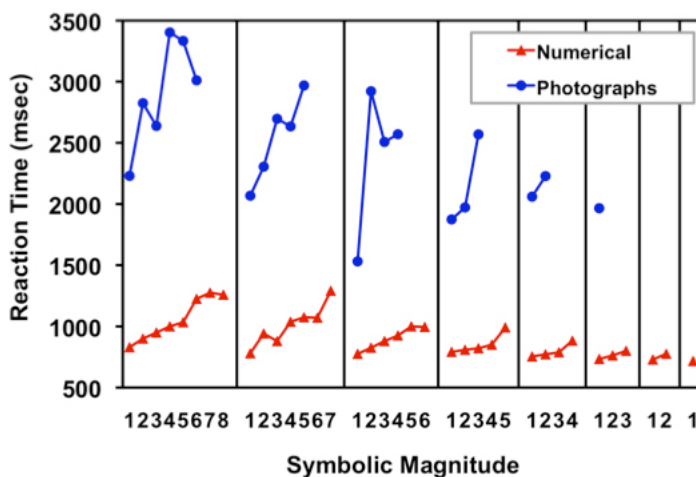


Figure 29. RTs of monkeys to 2-item subsets composed of numerical (Brannon & Terrace, 1998) and arbitrary (Terrace, Son, & Brannon, 2003) stimuli.

should it take 2-3 times as long to make a judgment about arbitrary stimuli than it does about numerical stimuli? Before addressing that question, we should note an interesting feature of the distance functions obtained with arbitrary stimuli that is absent from those obtained with numerical stimuli. In the case of arbitrary stimuli, RTs are often maximal to items from the middle of the list. Returning to the data shown in Figure 32, we saw that, at a distance of 1, RTs for the monkeys trained on 7-item arbitrary lists increased from 2,200 msec at the first position to 3,425 msec at position E, and then decreased to 2,900 msec at position H. By contrast, the absolute values of RT functions that were obtained from the same monkeys on numerical subsets increased monotonically as the numerical value of the first item of the subset increased. Figure 30 shows similar RT functions that were obtained from both human and non-human primates. In each instance, unimodal magnitude functions were obtained following training with arbitrary stimuli and monotonic magnitude functions following training with numerical stimuli.

The drop in RT for judgments of items from the end of an arbitrary list suggests that subjects scan a representation of those lists from both ends (Jou 1997). The asymmetry of those distance functions suggests that the rate of forward scanning for items from the beginning of an arbitrary list is faster than the rate of backward scanning of items from the end of the list. This is hardly surprising given the huge amount of practice that subjects have had while executing lists in a forward manner.

A dual scanning process should be expected with over-trained lists of arbitrary stimuli but not with over-trained lists of numerical stimuli. On arbitrary lists, the last item becomes more salient during overtraining. As a result, experienced subjects tend to scan a list inwards, starting from the first item, and backwards, starting from the last item. When, for example, a human subject is asked, which comes first, T or W, it is easier and faster to scan backwards from Z to W than forward from A to T. On a numerical list, the last item does not become salient because the value of the last item on a given numerical list is not necessarily the same as its value on another numerical list and because subjects rely exclusively on a uni-directional rule to execute such sequences.

The symbolic distance and magnitude functions we obtained from human and non-human primates, both for arbitrary and numerical stimuli, were similar to those reported in other studies. Indeed, when the functions shown in Figure 30 are combined with those reported in the literature, they coalesce into two neatly segregated and non-overlapping groups that can be defined by the absolute value of RTs: one for arbitrary stimuli (Figure 31, the other for numerical stimuli (Figure 32).

Mechanisms for Judgments of Relative Magnitude. The essentially linear increase in the RT of judgments of relative magnitude suggests a serial process, both in the case of arbitrary and numerical stimuli. However, the relatively long RT for judgments of the magnitude of arbitrary stimuli suggests that judgments of arbitrary stimuli may require an additional step that is not needed for relative judgments of numerical stimuli. On arbitrary lists, the subject first scans from the start point of its representation of an arbitrary list to locate the item in question. The more advanced the first item is on the original list, the longer it takes to determine its ordinal position. That process would account for the magnitude effects that were observed on both between- and within-list subsets.

A second process begins when the subject locates the ordinal position of each item. The smaller the ordinal distance between the first and the second items of the subset, the more difficult it is to discriminate their ordinal positions. That would be reflected in the number of iterations of the com-

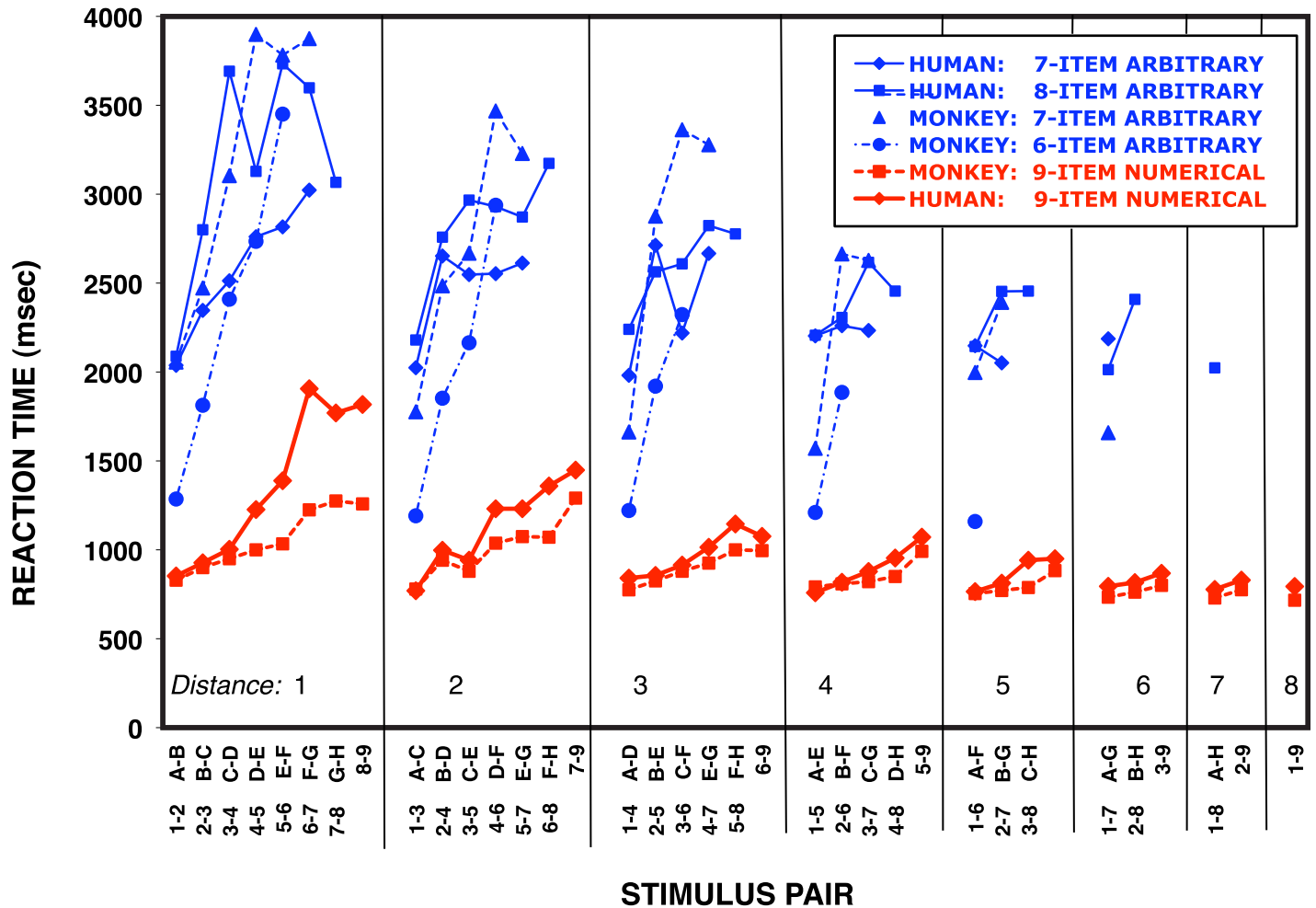


Figure 30. RTs of monkeys and humans to 2-item subsets composed of numerical and arbitrary stimuli containing 6 to 9 items. (Data are from experiments by Brannon & Terrace, 1998, 2002; Terrace, 2001; and Terrace, Son, & Brannon, 2003).

parison process that are needed to distinguish each item's location. Assuming that each iteration requires a fixed amount of time, the larger the number of iterations, the higher the value of the RT. Accordingly, the absolute value of the RT of a judgment of relative magnitude can be regarded as an indicator of the number of iterations needed to judge each combination of distance and magnitude.

Figure 33 depicts a serial process for judgments of the relative magnitude of arbitrary stimuli that are drawn from the same or different lists. When presented with a 2-item subset, the subject begins two scans, one for each item of the subset. In principle, the subject has enough information to choose the stimulus to which it should respond when it finds a match for one of the subset items. However, the distance and magnitude effects we observed suggest that the subject locates both items on the lists from which they were selected before deciding to which item it should respond first. There are many examples of completing a search that, logically, isn't necessary, e.g., Sternberg (1975).

For subsets composed of arbitrary items, we assume that

subjects must first retrieve ordinal information about each item from long-term memory (LTM). It then has to convert that information to its analog value in working memory (WM) and, finally, as shown in Figure 34, it has to compare those analog values to determine the ordinal value of each list item. For subsets composed of numerical items, we assume that the analog value of each item can be obtained directly in WM. There, a rule for moving from one item to the next is applied. Assuming that the retrieval of an ascending or descending rule from WM would take less time than retrieving the ordinal values of arbitrary stimuli from LTM, RTs of numerical judgments would be faster than RTs of arbitrary items.

Conversion of digital to analog representations of list items. When considering the SDE, it is easy to lose sight of the fact that subjects cannot base their judgments of relative magnitude on the physical properties of the stimuli they are shown. It has therefore been argued that discrete symbols such as Arabic numerals, letters and verbal list items must be converted into analog ordinal representations. To decide the order of symbolic stimuli, the ordinal positions of those

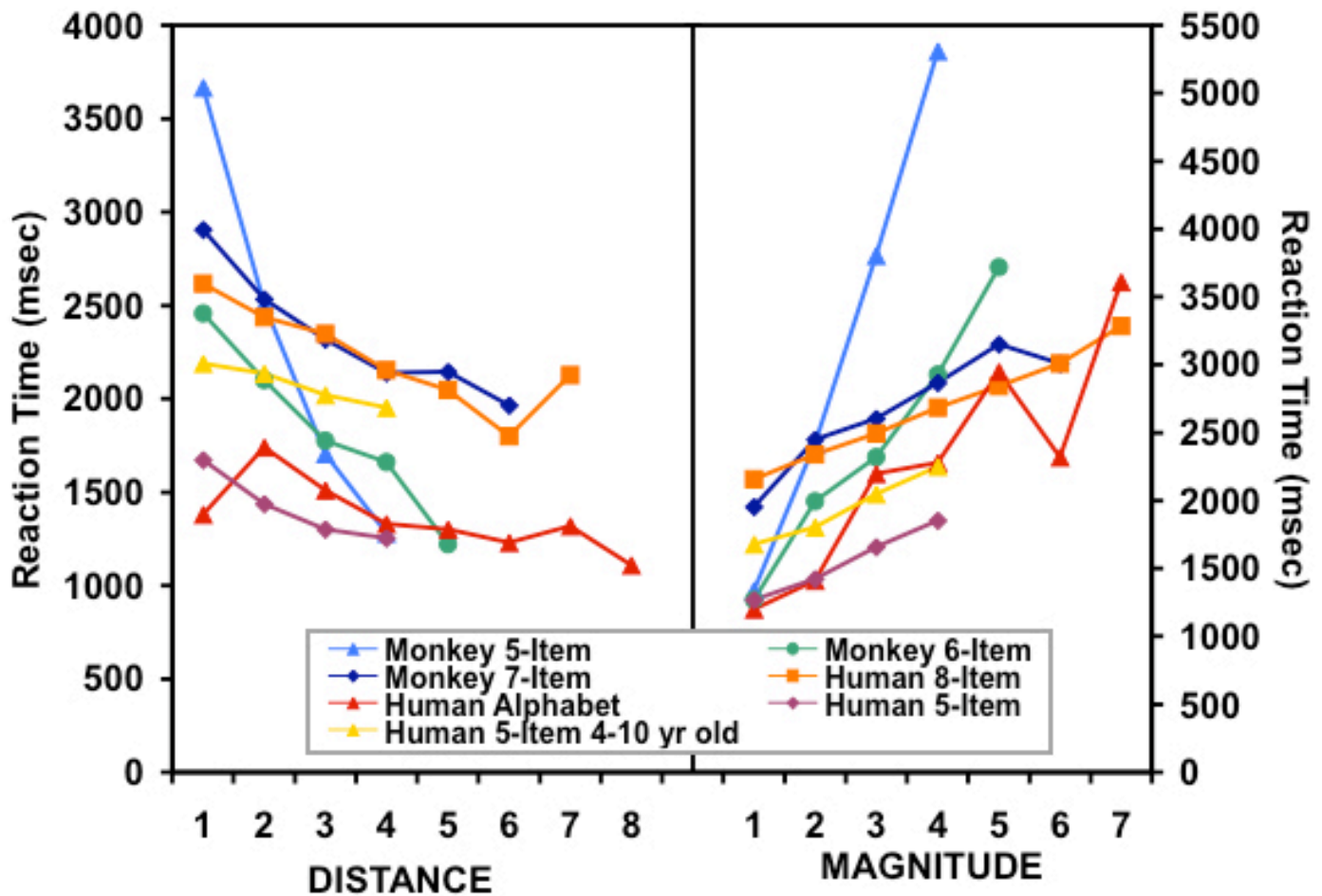


Figure 31. Distance and magnitude functions obtained from human subjects and monkeys. These functions were obtained on subtest tests composed of numerical and arbitrary items. All of the non-human primates were first trained using the simultaneous chaining paradigm to learn lists of arbitrary or numerical stimuli. They were then tested with 2-item subsets composed of items from the lists on which they were trained. The data shown are the mean of the RTs of correct responses to the first item of each subset. Also shown are distance and magnitude functions obtained from human subjects who were tested on their ordinal knowledge of letters of the alphabet or with arbitrary or numerical stimuli similar to those used in the experiments with non-human primates (Terrace, 2005). 1 = Brannon and Terrace (1998); 2 = Brannon and Terrace (2002); 3 = Murofushi (1997); 4 = D'Amato and Colombo (1988); 5 = Terrace, (2001b); 6 = Terrace et al. (2003); 7 = Terrace (2001a).

representations are then compared. The same is true for arbitrary stimuli (photographs) that have been associated with an ordinal position by virtue of training on the simultaneous chaining paradigm. Figure 35 shows how SDEs can be generated from nominal and ordinal stimuli. Each curve represents a response distribution of RTs to a particular stimulus. The width of each response curve increases as the magnitude (numerosity) or the ordinal position of an arbitrary stimulus increases. As a consequence, response curves overlap increasingly with increasing magnitude. The distance effect occurs because the distributions of items that are closer together overlap more than when items are farther apart.

Neural basis of numerosity discriminations. Similar distributions have been obtained from monkeys of the firing rate of single cells of the intraparietal sulcus (IPS) in frontal

cortex in response to different numerosities. Figure 36 show distributions obtained by Nieder and Miller (2004) of the output of single cells in the IPS while monkeys responded on a match/non-match task using arrays of geometric stimuli similar to those used by Brannon and Terrace (1998). Single unit recordings in the IPS indicated preferential tuning curves for the values 1-5. The width of the tuning curves increased logarithmically as the magnitude of the sample increased, thereby conforming to Weber's law, and creating an inherent order of increasingly overlapping tuning curves (Nieder, 2005). As is the case with behavioral studies of the symbolic distance and magnitude effects for numerical stimuli, these characteristics of numerical tuning curves can account for symbolic distance and magnitude effects. The concept of tuning derives from the specific responsiveness of different tuning forks to particular audio frequencies. In

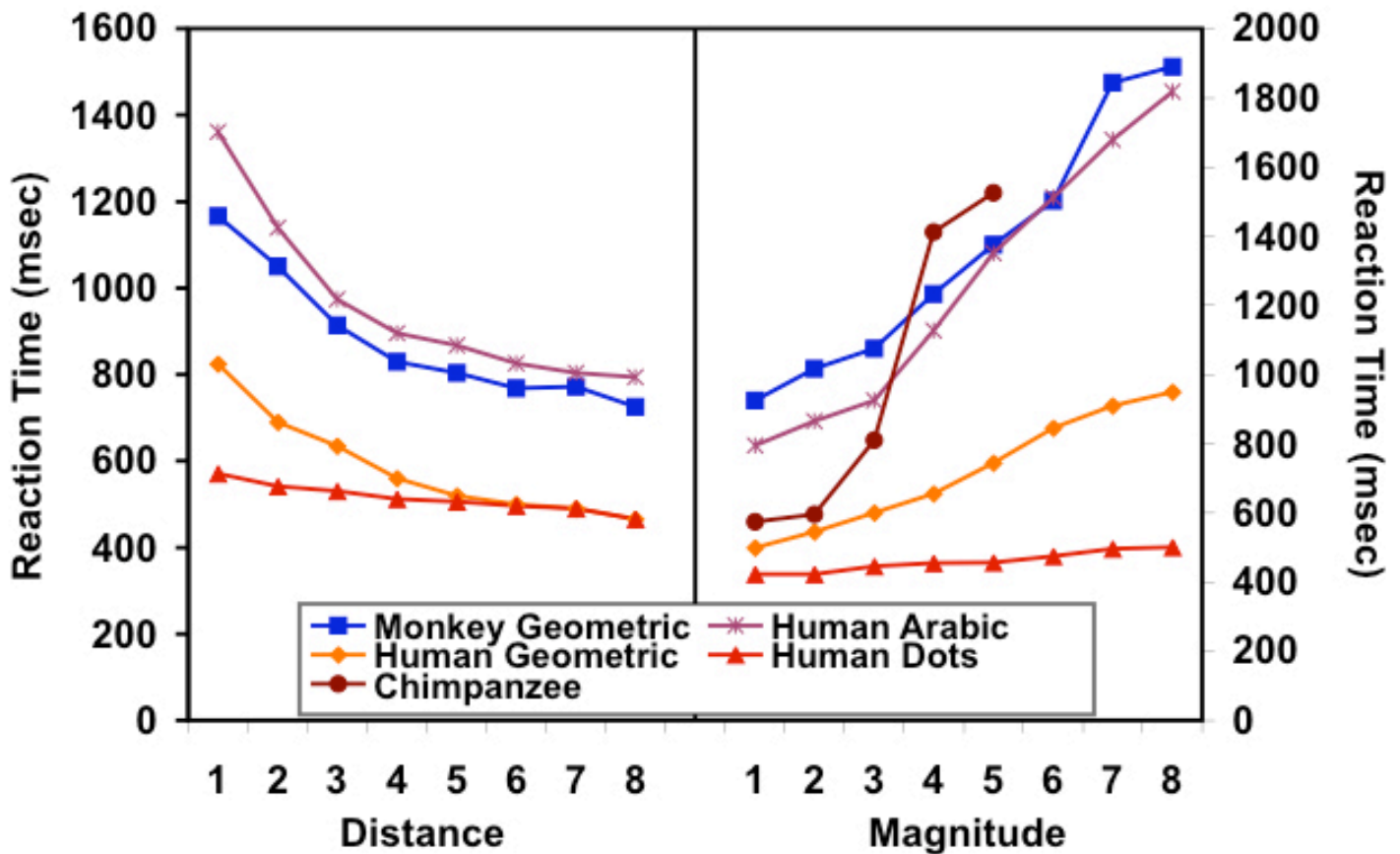


Figure 32. Distance and magnitude functions obtained from human subjects and monkeys. These functions were obtained on subtest tests composed of numerical and arbitrary items. All of the non-human primates were first trained to learn lists of arbitrary or numerical stimuli by the simultaneous chaining paradigm. They were then tested with 2-item subsets composed of items from a given list. The data shown are the median RTs of correct responses to the first item of each subset. Also shown are distance and magnitude functions obtained from human subjects who were tested on their ordinal knowledge of Arabic numbers or with arbitrary or numerical stimuli similar to those used in the experiments with non-human primates (Terrace, 2005). 8 = Hamilton and Sanford (1978); Brannon and Terrace (2002); 9 = Colombo and Frost (2001); Gulya and Colombo (2004); 10 = Gulya and Colombo (2004) 11 = Brannon and Terrace (2002); 12 = Buckley and Gillman (1974); 13 = Moyer and Landauer (1967); Buckley and Gillman (1974); 14 = Chalmers and McGonigle (1984).

this instance, tuning refers to the relative responsiveness of particular cells to different numerosities. The numerosity that evokes the largest response is the value assigned to a particular cell.

Transitive Inference. Symbolic distance and magnitude effects have been obtained in a wide variety of experiments on TI that are indistinguishable from those obtained following training on the simultaneous chaining paradigm. Some of these experiments used species whose serial abilities have not been previously investigated, e.g., lemurs and mongoose and cebus apella (McGonigle and Chalmers, 1977). Some examples are shown in Figures 37-38.

Unlike simultaneous chains, transitive inference (TI) has a long history in philosophy and mathematics before psychologists used it as a tool for studying reason. For example, a philosopher might ask a question in syllogistic form: if all men are mortal and Socrates is a man, is Socrates mortal?

A mathematician would use the same logic to ask a question such as, if $A > B$ and $B > C$, what is the relationship between A and C? Following Piaget, TI has been widely used to study the development of reasoning in children. The discovery that monkeys can use TI to solve problems based on relative order (McGonigle and Chalmers 1977), came as a surprise because animals were supposed to be incapable of using reason. More surprising was McGonigle and Chalmers' data from a subset test that they administered after subjects satisfied an accuracy criterion for training on adjacent pairs. As shown in Figure 37, McGonigle and Chalmers obtained a symbolic distance effect similar to those we've seen earlier in Figure 22. As far as I'm aware, this is the first evidence that an SME could be obtained from an animal. More recently, comparative psychologists have provided additional evidence of TI in animals in experiments on pigeons, corvids, rats and goldfish. Because the best evidence for TI comes from experiments on non-human primates, I will limit my

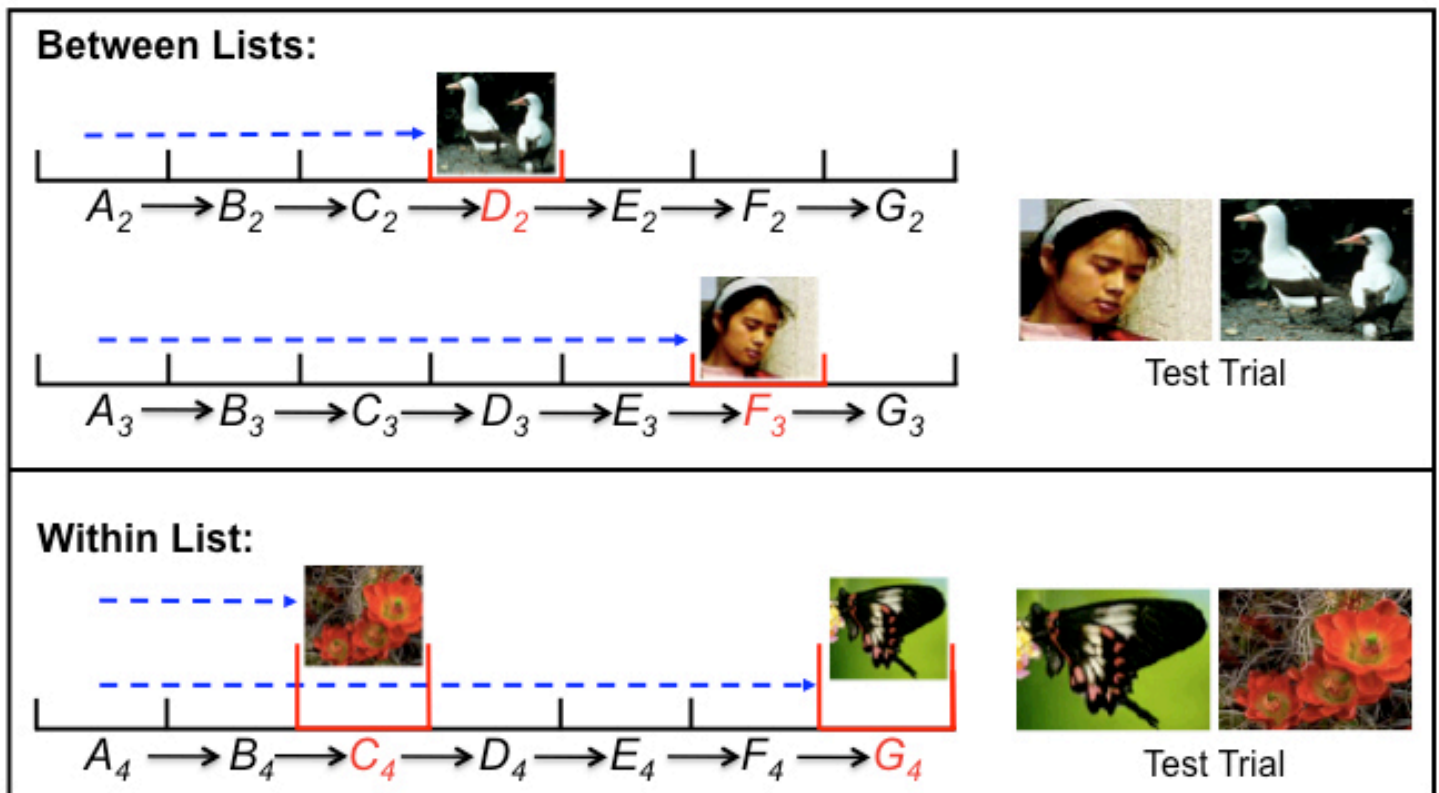


Figure 33. A mechanism for judgments of the relative magnitude of arbitrary stimuli that are drawn from the same or different lists. See text for additional information.

discussion of TI to that species.

The most impressive demonstration of TI in monkeys comes from in a recent experiment by Treichler, Raghanti et al. (2003). First, Treichler, et al. trained monkeys to respond accurately to adjacent pairs composed of 5 arbitrary geometric stimuli, $A > B > C > D > E$, by rewarding his subjects when they chose the larger item of a particular pair. In the examples that follow, the starred item is the one to which subjects must respond to obtain a reward. $A > B^*$, $B > C^*$, $C > D^*$ and $D > E^*$. He then trained 2 additional 5-item lists: $F > G^*$, $G > H^*$, $H > I^*$ $I > J^*$, and $K > L^*$, $L > M^*$, $M > N^*$ $N > O^*$ and $O > P^*$, and two linkage pairs, $E > F^*$ and $J > K^*$. The first pair, $E > F^*$, was inserted at the end of the first 5-item list; the second at the end of the second list. To show that the monkeys linked the 3 lists into a unitary 15-item list $A > B^*$, $B > C^*$, $C > D^*$, $D > E^*$, $E > F^*$, $F > G^*$, $G > H^*$, $H > I^*$, $I > J^*$, $J > K^*$, $K > L^*$, $L > M^*$, $M > N^*$, and $N > O^*$, Treichler tested his monkeys with a variety of pairings between non-adjacent items that spanned the 15-item list, e.g., B-M, G-K, J-N, etc. The monkeys' responses were just as accurate to non-adjacent pairs that were drawn from different lists as they were to non-adjacent pairs drawn from the same list. What's more, as shown in Figure 39, subjects' accuracy on the subset test increased with inter-item distance across all 3 lists and yielded the longest SDE to date.

Treichler's experiment is of interest for two reasons. The obvious reason is the sheer length of the lists the monkeys learned. It would, therefore, be of interest to determine if each 5-item list functioned as a chunk and whether Treichler's experiment could be described as the linking of 3 chunks. The other reason is the similarities (or difference) between representations of lists trained by the simultaneous chaining and the TI paradigms. There are at least 2 ways to address that question. The first is to compare the accuracies and RTs of monkeys trained by each paradigm in response to 2-item subsets and to then compare the similarity of symbolic distance and magnitude effects following each type of training. The second is to see if there is any transfer from knowledge of serial order, as first established by the TI paradigm, to knowledge of serial order that was obtained by learning a list trained by the simultaneous chaining paradigm, using the same list items, and vice versa. Given that the items used during each type of training are the same, familiarity with list items should not influence performance on each type of transfer test.

Other Properties of Ordinal Representations. We have seen that monkeys and humans, but not pigeons, can learn arbitrary and numerical lists, as trained by the simultaneous chaining paradigm. Striking similarities in distance and magnitude effects, obtained from monkeys and humans, suggest that both species form ordinal representations of

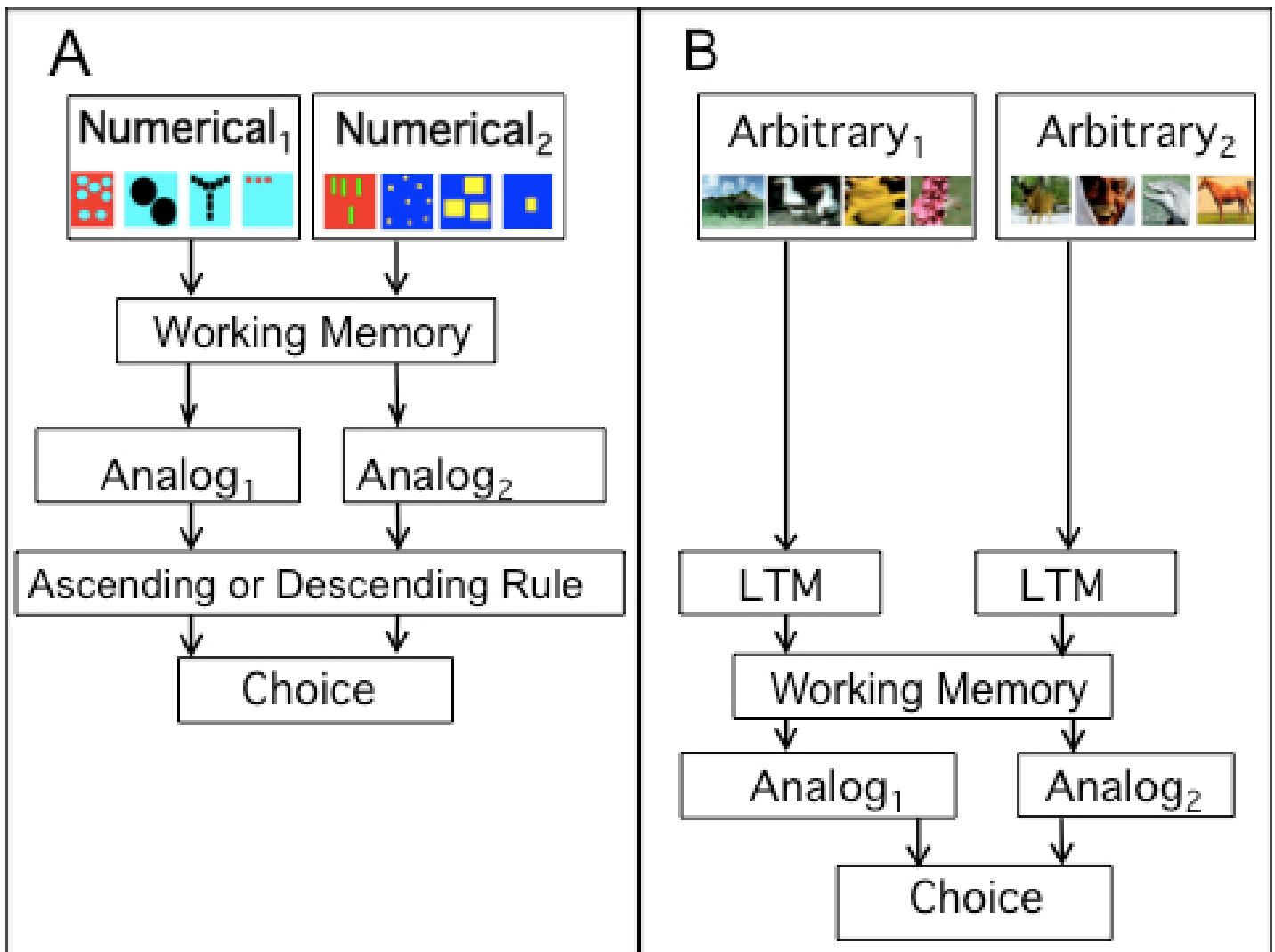


Figure 34. Processing of information derived from numerical and arbitrary stimuli during a 2-item subset test. See text for additional information.

the arbitrary and numerical sequences they learn to execute. In this section, we will describe a number of experiments whose common goal was to define properties of such representations.

Elasticity of Numerical Representations. Numerical and arbitrary continua differ in that magnitude is an inherent dimension. Thus, magnitude conveys both discrete information (which is stimulus A?) and ordinal information (does stimulus A come before or after stimulus B?). In a recent experiment, we investigated the discrete and continuous properties of numerical representation in monkeys using the matching-to-sample (MTS) paradigm with geometric stimuli, each composed of n elements. Research was conducted in three stages: identity MTS training, conceptual MTS training, and manipulation of stimulus range.

During Stage I (physically identical sample and test stimuli), two subjects were trained to match the numerosities 1-4

on the basis of physical similarity. In each instance, the correct choice was an exact replica of the sample stimulus. During Stage II (conceptual MTS training), the correct choice was conceptually identical to the sample but physically dissimilar. Examples of the stimuli used during Stages I and II are shown in Figure 40A. Monkeys were first tested with stimuli whose values ranged from 1-9 and, subsequently, from 1-15. For both ranges, there was a robust end effect, with higher accuracy and faster RTs at each end of the stimulus continuum than in the middle. As can be seen in Figure 40 B, the end effect was reflected in U-shaped accuracy functions. We also obtained inverted U-shaped functions for RTs (not shown). During Stage III, the range of the sample and the test stimuli were varied further. Initially, the range was 1-11. Both subjects were then tested with the ranges 3-13 and 5-15 (counterbalanced across subjects).

While response functions for all three ranges were qualitatively similar, the magnitude of each stimulus range ex-

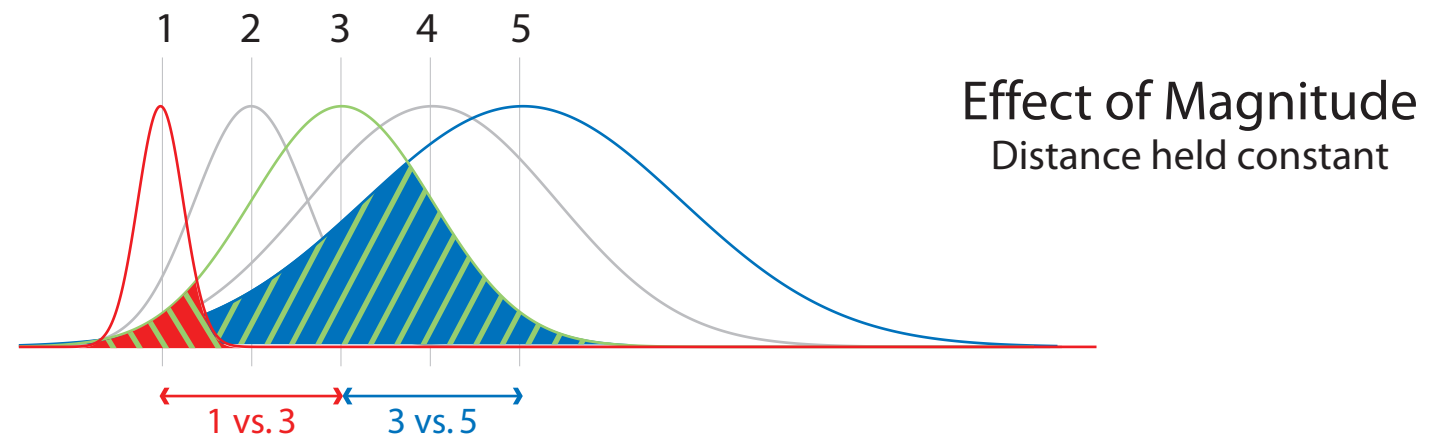
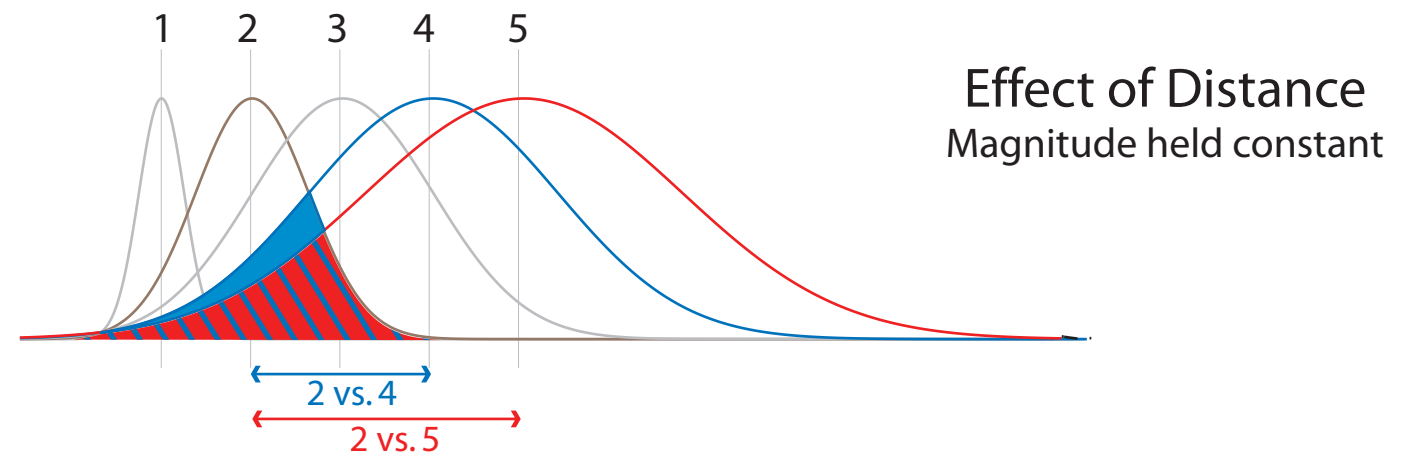


Figure 35. Comparisons of analog values of arbitrary and numerical stimuli that give rise to magnitude (upper panel) distance (lower panel) effects. The variance of judgments of numerosity and order increases as magnitude of ordinal position and numerical quantity increases. With magnitude held constant, larger distances are discriminated more readily in WM than smaller distances (upper panel). Similarly, with distance held constant, larger magnitudes are less discriminable than smaller magnitudes (lower panel).

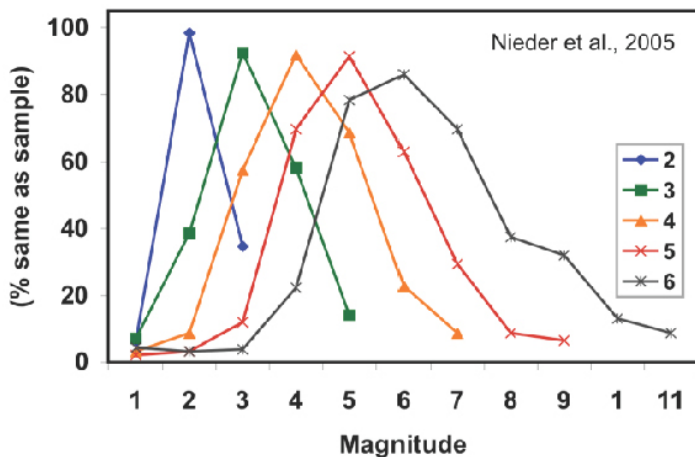


Figure 36. These functions are the averaged single-cell numerosity-tuning functions (from the PFC) (Nieder and Miller, 2005). The legend and the numerals within each function show the sample numerosity at which each curve was derived.

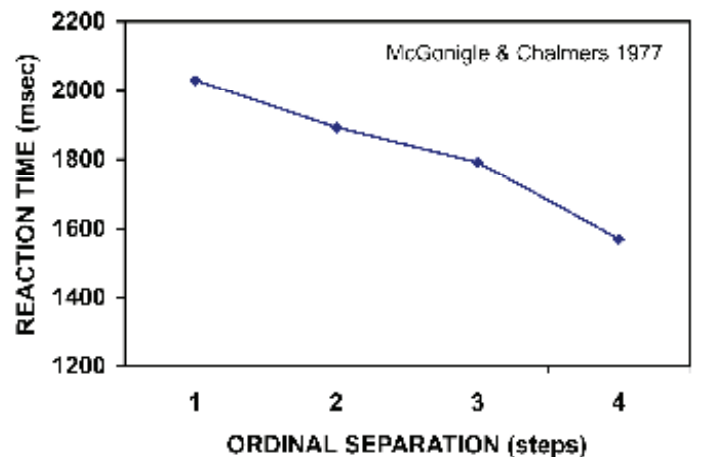


Figure 37. Magnitude function obtained from *cebus apella* after on a 5-item transitive inference task.

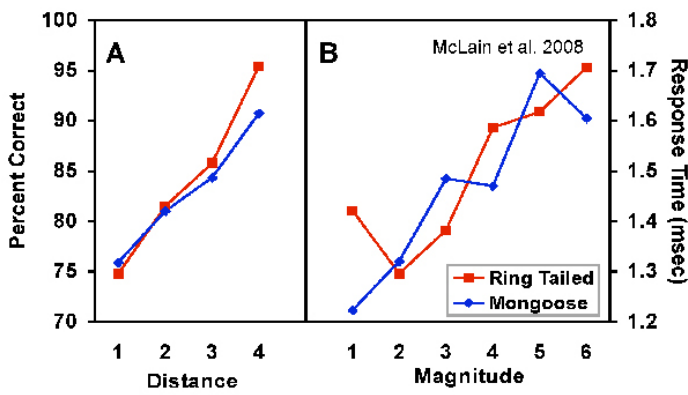


Figure 38. (A) Mean accuracy of lemurs as a function of symbolic distance. (B) Mean RTs to adjacent pairs. The higher-ranked stimulus in each adjacent pair decreases from left to right along the abscissa (MacLean et al., 2008).

erted a strong influence on accuracy, RTs to the sample and RTs to the test stimuli. Analysis of values shared by each range (5-11) indicated a significant effect of stimulus range, regardless of whether the entire range (1-11, 3-13) or only the shared values (5-11) were compared (Figure 41 A). Additionally, the duration of RTs to the sample differed significantly as a function of stimulus range, indicating that the representation of each stimulus range differed, even in the absence of the test stimuli (Figure 41 B). These results also indicate that, while stimulus magnitude does significantly affect numerical representation, context, as reflected in rela-

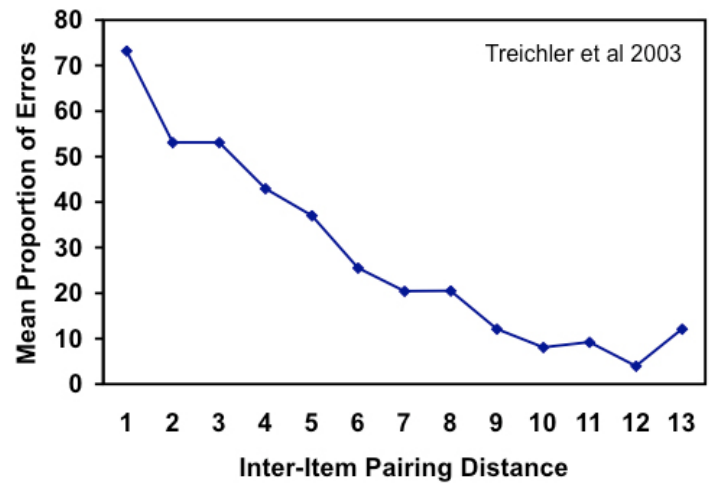
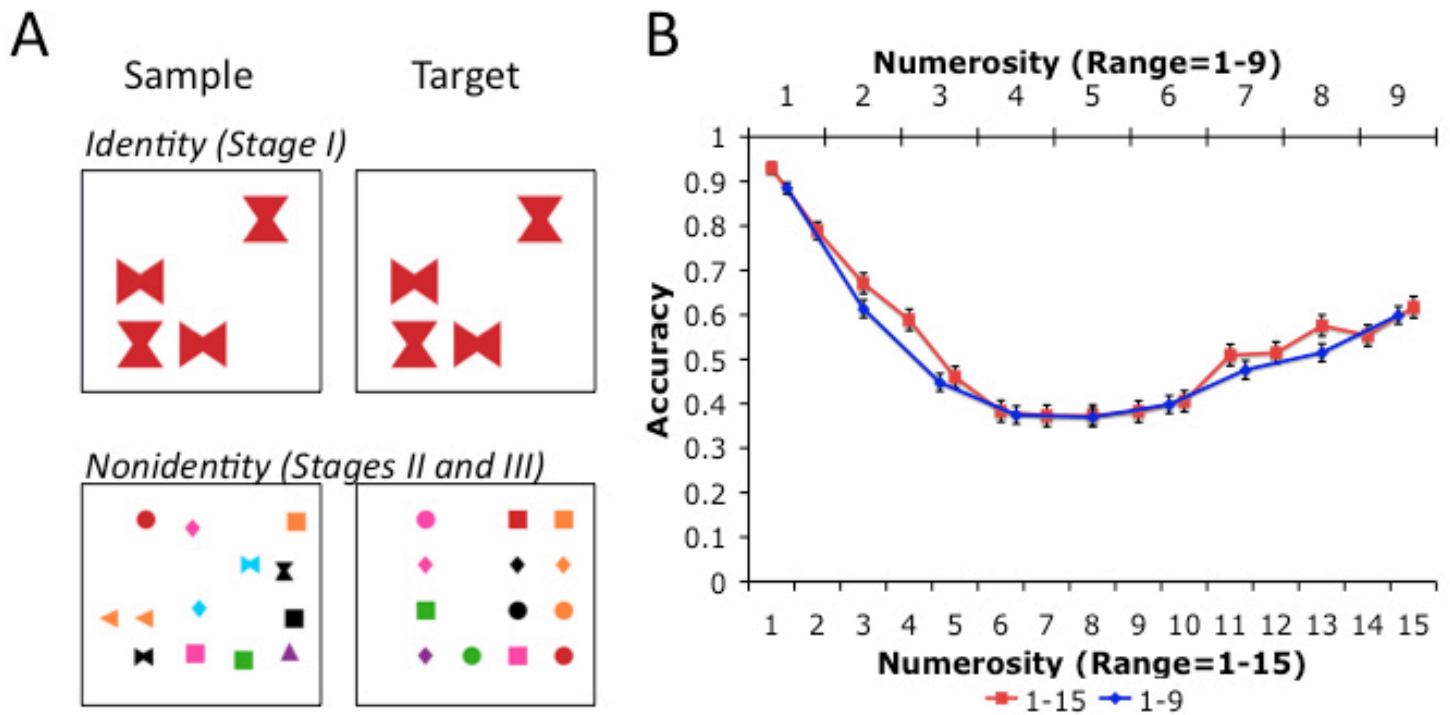


Figure 39. Mean proportion of errors at each possible inter-item pairing distance (relative to the 15-item list) for linked conditions for the entire 15-day test (Treichler et al., 2003).

tive continuum position and stimulus range, is also a significant factor.

The effect of stimulus range suggests that monkeys form expectations for a range of stimulus values and that a stimulus' relative position within that range influences the accuracy and the RT of responses. These list-like aspects of numerical continua, along with symbolic distance effects and the effect of range magnitude, suggest that, even in the absence of ordinal training, monkeys represent numerical con-



Matching to Numerical Stimuli

Figure 40. (A) Examples of sample and test stimuli; (B) Accuracy for ranges 1-9 and 1-15 (Stage II). Bars show standard errors.

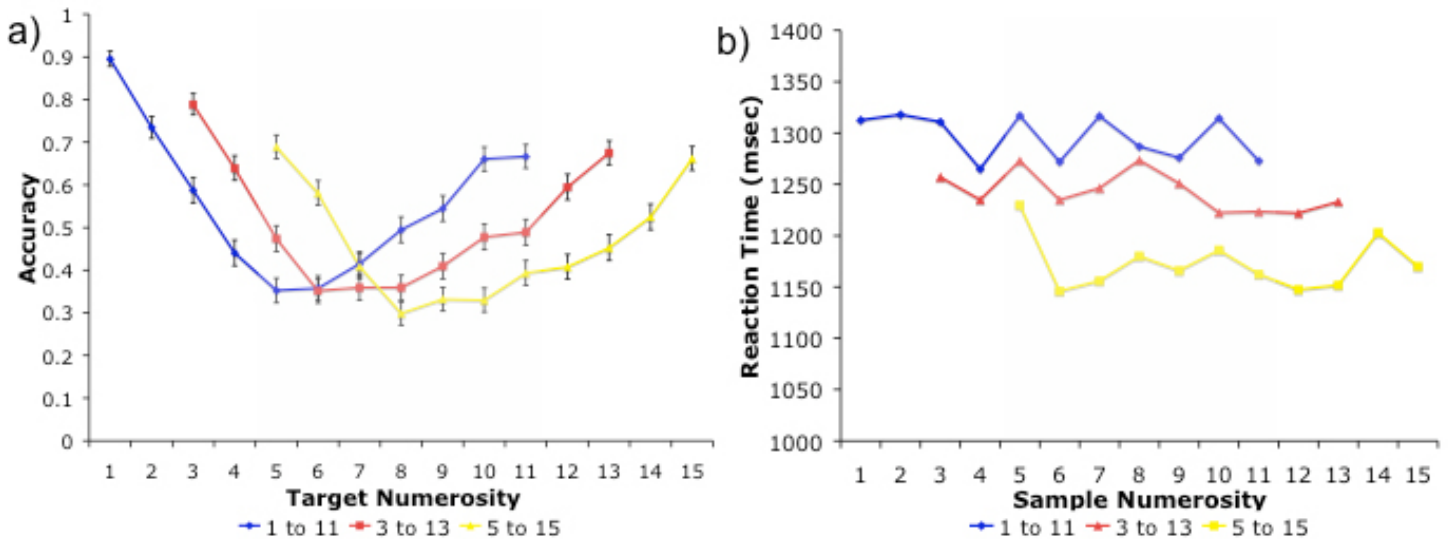


Figure 41. The gray highlighted section shows the numerosities that are common to all three stimulus ranges (Stage III). (A) Accuracy for 1-11, 3-13, and 5-15. Bars reflect standard errors. (B) Reaction times to the sample for 1-11, 3,13, and 5-15. Error bars are not shown because a high degree of variability.

tinua continuously and that they integrate discrete ordinal information about numerosity with an overall representation of the stimulus range.

Relative vs. Absolute Positional Codes. Suppose that you were the manager of a baseball team and you received an offer from another manager to trade your top 3 hitters for the top 3 hitters of his team. Is this a fair trade and, if not, why? You can't answer that question unless you know the actual batting averages of the players involved. The batting averages of the top 3 hitters on your team, say .325, .298 & .349 may be higher than the averages of the top 3 batters of the other team, .276, .246 & .315. The moral is that the absolute value of the top ranks of two different distributions can obscure the relative values that are represented by those ranks and it is the relevant values that should be the basis of your decision.

In a recent experiment we tested monkeys on a similar problem (Merritt & Terrace, 2004). We first trained 2 monkeys on a 5-item TI problem, ABCDE. On that list C occupies the middle (or 3rd) position. We then trained the same monkeys to learn a 9-item TI problem, ABCDEFGHI. On that list, E occupies the middle position. Association models would predict an absolute rather than a relative spacing of items along an ordinal continuum ("1st", "2nd", or "3rd", etc.). A relative model predicts relative spacing of list items ("beginning", "middle", and "end"). The aim of this experiment was to show that ordinal information is based on the relative values of list items trained by the transitive inference paradigm, as opposed to an absolute code model which predicts uniform spacing ("1st", "2nd", or "3rd", etc.).

Two monkeys were trained on two lists, one containing

5 items, the other, 9 items. On both lists (which are shown schematically in Figure 42), subjects were trained by the TI paradigm on adjacent pairs. On a transitive inference paradigm a subject is trained on adjacent items of some list. For example, on a 6 item list, ABCDEF, subjects would be trained on the following pairs, AB*, BC*, CD*, DE* and EF*. The critical test following this type of training is, how would subjects respond to non-adjacent items, e.g., DB, CE, CF, etc? In this experiment subjects were trained to an accuracy criterion for each of 80% correctly completed trials on adjacent items before non-adjacent items were introduced. For a 5-item list, the pairs were A-B*, B-C*, C-D*, and D-E*, where the asterisk shows the item that the subject had to select for a reward. For a 9-item list, the pairs were A-B*, B-C*, C-D*, D-E*, E-F*, and F-G*. After a subject satisfied the accuracy criterion on each list, they were tested on between-list pairs to determine whether their choices were based on an item's relative or absolute position. Of particular interest were items that occupied identical absolute positions on each list, where absolute and relative encoding predict opposite choices. Consider, for example, the 2 lists shown in Figure 42. Suppose that the pair D5E9, is presented which depicts a 9-item list and absolute (Abs) and relative (Rel) representations of a 5-item list. If a subject relied on an absolute representation of ordinal position, it would choose E over D ($E9_{Abs} > D5_{Abs}$). If it relied on relative position, it would choose D over E ($D5_{Rel} > E9_{Rel}$). Similar to the results of comparable experiments with human subjects, we found that subjects relied on an item's relative position when judging ordinality. This result is important for two reasons. First, it provides additional evidence that monkeys represent ordinal position on a non-associative ordinal continuum. Second, it suggests that a monkey can determine an

Absolute Representation:

| | | | | | | | | |
|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| A ₉ | B ₉ | C ₉ | D ₉ | E ₉ | F ₉ | G ₉ | H ₉ | I ₉ |
| A ₅ | B ₅ | C ₅ | D ₅ | E ₅ | | | | |

Relative Representation:

| | | | | | | | | |
|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| A ₉ | B ₉ | C ₉ | D ₉ | E ₉ | F ₉ | G ₉ | H ₉ | I ₉ |
| A ₅ | | B ₅ | | C ₅ | | D ₅ | | E ₅ |

Figure 42. Design of experiment to determine if monkeys rely on absolute vs. relative values of list items when comparing numerical values from different lists. See text for additional details.

item's ordinal position outside of the context on which it was trained.

Directionality of Ordinal Rules. In our original experiment on the numerical ability of monkeys (Brannon and Terrace, 1998) one subject was trained to respond to exemplars of the numerosities 1-4 in a descending (4-3-2-1) as opposed to an ascending (1-2-3-4) order. That subject's accuracy on test pairs composed of the novel numerosities 5-9 was significantly lower than the accuracy of the subjects trained on the ascending series.

Because our subjects had equal amounts of training on ascending and descending sequences, there was no reason to believe that they would find it more difficult to "count backwards" than forward, as a child might. However, there was one difference in the application of the ascending and descending rules that the subjects learned when they were tested on the novel numerosities 5-9. Following training on the descending rule, subjects had to reverse direction when responding to those novel numerosities, whereas a subject that was trained on an ascending rule did not. That difference prompted us to train 2 monkeys to learn an ascending or a descending sequence composed of the numerosities 4, 5 and 6 (4-5-6 & 6-5-4). As shown in Figure 43 this allowed us to test subjects with novel stimuli whose numerical values were both larger and smaller than the values on which they were originally trained (Brannon, Cantlon et al., 2006).

Following 4-5-6 or 6-5-4 training, each subject was tested on all pairs of numerical values that could be derived from the numerosities 1-9. As in our previous experiment (Brannon and Terrace, 1998), the pairs were categorized as Familiar-Familiar (FF), Familiar-Novels (FN) or Novel-Novels (NN). Curiously, there was no decrement in performance on NN pairs as compared to FF or FN pairs, as was the case in the pair-wise tests that followed 1-2-3-4 and 4-3-2-1 training. However, a more detailed analysis, in which the NN

pairs were sub-divided into 3 sub-categories, revealed that the average accuracies we obtained were misleading. As shown in Figure 44, the small category consisted of pairs in which both items were smaller than the training values (1-2, 1-3, or 2-3); the large category, pairs in which both items were larger than the training values (7-8, 7-9, 8-9), and a "span" category, pairs of one small and one large item (1-7, 1-8, 1-9, 2-7, 2-8, 2-9, 3-7, 3-8, 3-9). All 3 monkeys performed well when presented with pairs from the span category. However, as shown in Figure 44, a more complicated pattern emerged for the pairs composed exclusively of small and large values. The monkey trained to produce a 4-5-6 sequence performed well when tested with large novel values (7, 8 & 9) but not when tested with small novel values (1, 2 & 3). By contrast, the two monkeys trained to produce 6-5-4 sequences performed well when tested with small novel values but not when tested with large novel values. It appeared therefore that, during their initial numerical training, monkeys learned an abstract unidirectional rule, "respond to the larger (or smaller) value first." Accordingly, accuracy of responding for monkeys trained on the 6-5-4 sequence was higher to small novel values than to large novel pairs, and accuracy of the monkey trained on the 4-5-6 sequence was higher to large than to small novel pairs.

Neurological Mechanisms of Ordinal Knowledge. Typically, neuroscientific discoveries about behavioral mechanisms occur only after a firm behavioral foundation has been established about the mechanism in question. In recent years that principle has begun to bear fruit in neuroscientific investigations of serially organized behavior. We have already seen an example of how behavioral data on a monkey's ability to discriminate numerosity led to the discovery of single cells in the parietal cortex that are maximally responsive to particular numerosities (Neider, 2005).

Neural representations of ordinal sequences. A major

| | | | | | | | | |
|--------------------|---|---|-------------------|---|---|------------|---|---|
| Ebbinghaus: | | | | | | | | |
| Reversals | | | Ascending Rule → | | | Extensions | | |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| Poor | | | | | | Good | | |
| Lashley: | | | | | | | | |
| Extensions | | | Descending Rule ← | | | Reversals | | |
| 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 |
| Good | | | | | | Poor | | |

Figure 43. Design of experiment for evaluating subset test performance following training on ascending (4-5-6) and descending (6-5-4) training (Brannon, Cantlon and Terrace, 2006). See text for additional details.

theme of recent work on the serial organization is the type of representation used to organize serial information. In that connection, we have reviewed a considerable amount of evidence that shows that stimuli from different dimensions can be encoded as analog values and compared along a single dimension-independent continuum of order. Specifically, we have seen that RT and accuracy functions that give rise to SDEs and SMEs are quite similar across many different perceptual and psychological continua, as judged by humans and monkeys. These similarities suggest a common form of representation that mediates performance on numerical, psychophysical, serial, and transitive inference tasks. We have also seen that representations that are based on arbitrary stimuli, whose order has to be learned, are similar to representations of numerical stimuli that are inherently ordered.

Although evolutionary and biological considerations provide different rationales for characterizing ordinal representations as spatial, temporal, or some combination thereof, I think that the available evidence favors spatial representations. Temporal representations do not apply for two reasons. The first is the absence of any temporal constraints on the responses during the execution of a simultaneous chain. The second is that the variability of reaction and inter-response times is too large to account for the high degree of accuracy of responding we have observed on various serial tasks (Terrace, Son et al., 2003). By default, it seems reasonable to opt for a spatial representation, but that conclusion would rest on firmer ground if there were a methodology to vary space in the same manner that has been done in experiments on

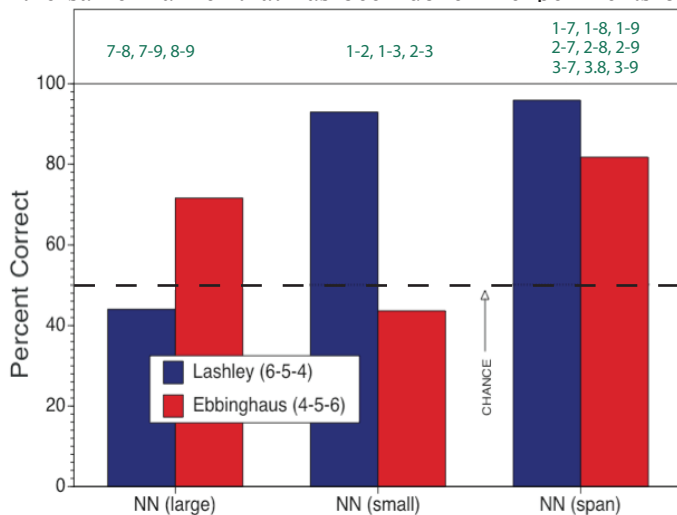


Figure 45. Performance on novel-novel numerosity pairs segregated as a function of whether the numerical values were, both small (1–2, 2–3, or 1–3), both large (7–8, 8–9, or 7–9), or whether they spanned the training values and involved one small (1, 2, or 3) and one large (7, 8, or 9) numerosity. Error bars reflect variance across test sessions. Values in parentheses indicate the rule on which each of the monkeys was trained.

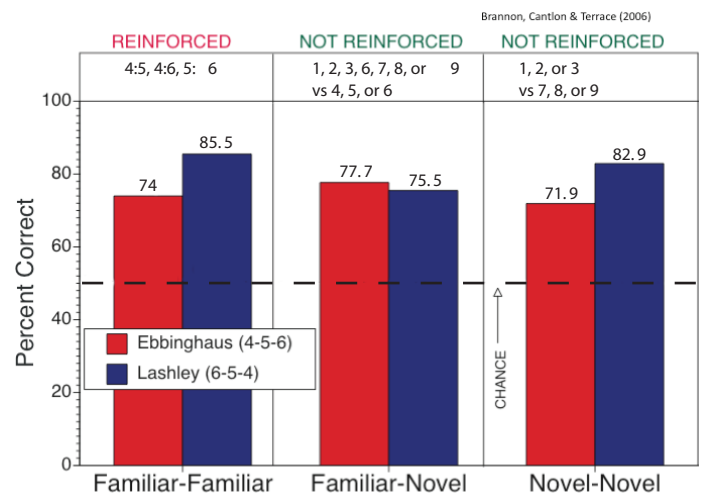


Figure 44. Performance on familiar-familiar, familiar-novel, and novel-novel numerosity pairs. Only familiar-familiar pairs were reinforced. Error bars reflect variance across test sessions. Values in parentheses indicate the rule on which each of the monkeys was trained.

timing (Gibbon & Church, 1900; Meck & Church, 1983). In such experiments, it is only necessary to vary the duration of a single stimulus to vary the temporal representation of that stimulus (Gibbon, 1990; Meck & Church, 1985). Since, as noted earlier, it is necessary to use at least 3 stimuli to compose a simultaneous chain, and because all of the items used to compose a simultaneous chain are presented simultaneously, the only way to manipulate a subject's spatial representation of those items is to vary the neural correlates of that representation, a feat that is beyond the grasp of current technology. It would also be desirable to provide an interval metric for spatial representations. The strongest case for an interval metric has been made in the case of numerical stimuli that, when plotted on a logarithmic scale, seem to follow Weber's law. However, that is the exception rather than the rule for the various continua we have used.

Marshuetz, et al. (2006) conducted a task that required human subjects to view sequences of five letters, and at test, to state if a probe letter occurred in the same order as it did during the sample. The authors found that activity of the left parietal cortex was inversely correlated with inter-item distance and concluded that order information was mediated in the parietal cortex by integrating them with magnitude codes.

Studies of the neural bases of numerical and spatial representation have also observed correlated activity in the parietal regions. Right parietal lesions may result in left hemispatial neglect, in which patients neglect items on the left side of a stimulus array. Because hemispatial neglect affects the perception of visual stimuli, patients performing a line bisection task displace the midpoint of the line toward

the right side. More remarkably, patients also displace the midpoint when asked to verbally bisect a numerical interval, suggesting a common locus of spatial and numerical representation in the parietal lobe (Dehaene and Changeux 1993; Zorzi, Priftis et al. 2006).

In contrast to spatial and psychophysical continua, ordinal relationships between lists of arbitrary stimuli must be learned. As a result we would expect different brain areas to be involved in learning about those stimuli than those used in inherently ordered tasks. Recent evidence from a variety of sources implicates the hippocampus as a primary site for learning ordinal information about arbitrary stimuli (Austin & Buckmeister, 2004; Dusek and Eichenbaum, 1997; Greene 2001; Heckers, Zalesak et al., 2004).

In one experiment, monkeys were trained on a transitive inference paradigm with each stimulus presented behind one of two food wells (Austin & Buckmeister, 2004). Together, the stimuli formed an ordered list, and, on any given trial, selection of the stimulus that occurred earlier in the sequence was rewarded. Selective damage to the entorhinal cortex, a major relay path between the hippocampus and the cortex, impaired performance on the novel BD comparison, while mainly sparing conditional associations between training pairs or AE comparisons to which the subject could respond correctly on the basis of its history of reward. The hippocampus appears to be involved in nonverbal transitive inference performance in humans as well. Heckers et al. (Heckers, Zalesak et al. 2004) compared blood oxygen level dependent signals obtained from subjects while they were performing a category inference task (not transitive) with stimuli obtained during the performance on a TI task. Only the hippocampus showed greater activation while subjects were making transitive inferences. It appears, therefore, that the hippocampus is uniquely engaged in making judgments of TI in humans as well as animals suggesting that the activation differential in the hippocampus is not due to a verbal strategy or the implementation of a conscious logical rule (Greene, Gross et al., 2006).

Taken together, these studies show that the hippocampus is required for comparing the ordinal position of items in the absence of specific item-item associations, that it doesn't require verbal ability or conscious awareness, and that it is common to both humans and animals. In a similar vein, it has been argued that the hippocampus appears to be the locus of learned hierarchical relationships between stimuli, whereas the parietal lobe is the locus of inherent spatial and numerical order discriminations (Dehaene, 1997).

In addition to the hippocampus, at least one study suggests that the parietal lobe is involved on tasks that require comparisons of arbitrary stimuli. Marshuetz et al. (2006)

conducted a task in which human subjects were required to view sequences of five letters. At test, they were required to state if a probe letter appeared in the same order as it did during the sample. The authors found left parietal activity to be inversely correlated with inter-item distance and concluded that order information is mediated in the parietal cortex by virtue of its integration with magnitude codes. This theory fits well with the hypothesis that magnitude information about the order of arbitrary list items is also mediated by parietal cortex.

Behavioral evidence that symbolic distance and magnitude effects that are generated by stimuli from different dimensions raises an interesting question. Are there separate systems for making ordinal judgments of nominal and ordinal stimuli from qualitatively different continua? Parsimony suggests that a central mechanism that compares the analog values of qualitatively different stimuli would be more economical than specific mechanisms, each devoted to a particular type of stimulus. On this view, ordinal information from a broad set of continua is formatted analogically by peripheral mechanisms or in working memory and is then evaluated in parietal cortex. Recent experiments on single cell recording from monkeys (Nieder and Miller 2004) and brain imaging in humans (Dehaene, Molko et al. 2004; Cantlon, Brannon et al. 2006) suggest that parietal cortex might be the sight of an "ordinal comparator". That would provide another example of the discovery of a brain mechanism that was based on a behavioral analysis of its features.

Conclusions

In this article, I discussed experiments that showed how the simultaneous chaining paradigm is an effective method for training sequences composed of arbitrary and numerical stimuli. Those experiments also provided evidence of an overlap in the mechanisms responsible for ordering arbitrary and numerical stimuli in monkeys and humans. Both monkeys and humans cannot only learn associations between adjacent items on arbitrary lists, but they also learn the relative ordinal position of all list items without any requirement to do so. For example, it was shown that monkeys were able to learn lists derived from items used on prior lists, immediately, so long as the items used to construct the derived lists maintained the item's original ordinal position. If their ordinal positions were changed, it took just as long to learn a derived list as list composed of novel items. In other experiments, it was shown that monkeys and humans were able to respond accurately to two-item subsets composed of adjacent and non-adjacent items from a particular list. Indeed,

they were able to do so, even when the subsets were composed of items from different lists.

Additional evidence of the spontaneous acquisition or ordinal knowledge was provided by the similarity in the slopes of distance and magnitude functions that were obtained from both monkeys and humans on two different serial tasks: transitive inference and simultaneous chaining. At the neurological level, there is evidence that mechanisms for ordering both numerical and arbitrary stimuli are located in the right parietal cortex.

It therefore seems reasonable to conclude that non-human primates have non-verbal mechanisms for learning lists composed of arbitrary items and for learning ascending and descending rules for ordering numerical stimuli. That helps to place in perspective the evolution of our abilities to count and to engage in conversation. If non-human primates lacked the capacity to learn numerical and arbitrary sequences without language, it is doubtful that humans would be able to order quantities according to their numerosities or to converse in utterances containing more than a single word.

References

- Atneave, F. (1955). Symmetry, information, and memory for patterns. *American Journal of Psychology*, 68, 209-222. doi:[10.2307/1418892](https://doi.org/10.2307/1418892).
- Brannon, E., Cantlon, J., & Terrace, H. (2006). The role of reference points in ordinal numerical comparisons by rhesus macaques (*Macaca mulatta*). *Journal of Experimental Psychology: Animal Behavior Processes*, 32(2), 120-134. doi:[10.1037/0097-7403.32.2.120](https://doi.org/10.1037/0097-7403.32.2.120).
- Brannon, E. and H. S. Terrace (2010). Analog Representation of the Numerosities 1-9 by Rhesus Macaques. *Journal of Experimental Psychology: Animal Behavior Processes* (under review).
- Brannon, E., & Terrace, H. (1998). Ordering of the numerosities 1 to 9 by monkeys. *Science*, 282(5389), 746-749. doi:[10.1126/science.282.5389.746](https://doi.org/10.1126/science.282.5389.746).
- Brannon, E., & Terrace, H. (2000). Representation of the numerosities 1-9 by rhesus macaques (*Macaca mulatta*). *Journal of Experimental Psychology: Animal Behavior Processes*, 26(1), 31-49. doi:[10.1037/0097-7403.26.1.31](https://doi.org/10.1037/0097-7403.26.1.31).
- Brannon, E. M. and H. S. Terrace (2002). The evolution and ontogeny of ordinal numerical ability. *The Cognitive Animal: Empirical and Theoretical Perspectives on Animal Cognition*. M. Bekoff, C. Allen and G. M. Burghardt. Cambridge, MIT Press: 197-204.
- Cantlon, Brannon, et al. (2006). Functional imaging of numerical processing in adults and 4-y-old children. *PLoS Biology* 4(5): 844-854. doi:[10.1371/journal.pbio.0040125](https://doi.org/10.1371/journal.pbio.0040125).
- Chen, S., Swartz, K., & Terrace, H. (1997). Knowledge of the ordinal position of list items in rhesus monkeys. *Psychological Science*, 8(2), 80-86. doi:[10.1111/j.1467-9280.1997.tb00687.x](https://doi.org/10.1111/j.1467-9280.1997.tb00687.x)
- D'Amato, M., & Colombo, M. (1988). Representation of serial order in monkeys (*Cebus apella*). *Journal of Experimental Psychology: Animal Behavior Processes*, 14(2), 131-139. doi:[10.1037/0097-7403.14.2.131](https://doi.org/10.1037/0097-7403.14.2.131).
- D'Amato, M. R. and M. Colombo (1990). The symbolic distance effect in monkeys (*Cebus apella*). *Learning and Behavior*, 18(2): 133-140.
- Davis, H., & Pérusse, R. (1988). Numerical competence in animals: Definitional issues, current evidence, and a new research agenda. *Behavioral and Brain Sciences*, 11(4), 561-615.
- Dehaene, S. (1997). *The number sense: How the mind creates mathematics*. New York, Oxford University Press.
- Dehaene, S., & Changeux, J. (1993). Development of elementary numerical abilities: A neuronal model. *Journal of Cognitive Neuroscience*, 5(4), 390-407. doi:[10.1162/jocn.1993.5.4.390](https://doi.org/10.1162/jocn.1993.5.4.390).
- Dehaene, S., & Cohen, L. (1997). Cerebral pathways for calculation: Double dissociation between rote verbal and quantitative knowledge of arithmetic. *Cortex*, 33(2), 219-250. doi:[10.1016/S0010-9452\(08\)70002-9](https://doi.org/10.1016/S0010-9452(08)70002-9).
- Dehaene, S., Molko, N., Cohen, L., & Wilson, A. (2004). Arithmetic and the brain. *Current Opinion in Neurobiology*, 14(2), 218-224. doi:[10.1016/j.conb.2004.03.008](https://doi.org/10.1016/j.conb.2004.03.008).
- Dusek, J. A. and H. Eichenbaum (1997). The hippocampus and memory for orderly stimulus relations. *Proceedings of the National Academy of Sciences* 94(14): 7109-7114.
- Ebbinghaus, H. (1964). Rapidity of learning series of syllables as a function of their length. *Memory: A Contribution to Experimental Psychology*. New York, Dover: 46-51.
- Ebenholtz, S. (1963). Position mediated transfer between serial learning and a spatial discrimination task. *Journal of Experimental Psychology*, 65(6), 603-608. doi:[10.1037/h0040458](https://doi.org/10.1037/h0040458).
- Gallistel, C., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition*, 44(1-2), 43-74. doi:[10.1016/0010-0277\(92\)90050-R](https://doi.org/10.1016/0010-0277(92)90050-R).
- Gallistel, C., & Gelman, R. (2000). Non-verbal numerical cognition: From reals to integers. *Trends in Cognitive Sciences*, 4(2), 59-65. doi:[10.1016/S1364-6613\(99\)01424-2](https://doi.org/10.1016/S1364-6613(99)01424-2).
- Gibbon, J., & Church, R. (1990). Representation of time. *Cognition*, 37(1-2), 23-54. doi:[10.1016/0010-0277\(90\)90017-E](https://doi.org/10.1016/0010-0277(90)90017-E).
- Gelman, R. G., R.C. (1978). *The Child's Concept of Number*. Cambridge, Harvard University Press.
- Greene, A. J., W. L. Gross, et al. (2006). An fMRI analysis of the human hippocampus: Inference, context, and task awareness. *Journal of Cognitive Neuroscience*, 18(7), 1156-1173. doi: [10.1162/jocn.2006.18.7.1156](https://doi.org/10.1162/jocn.2006.18.7.1156)
- Greene, A. J., Spellman, B. A., Dusek, J. A., Eichenbaum,

- H. B., & Levy, W. B. (2001). Relational learning with and without awareness: transitive inference using nonverbal stimuli in humans. *Memory and Cognition*, *29*(6), 893-902.
- Hamilton, J., & Sanford, A. (1978). The symbolic distance effect for alphabetic order judgements: A subjective report and reaction time analysis. *The Quarterly Journal of Experimental Psychology*, *30*(1), 33-41.
- Harlow, H. (1949). The formation of learning sets. *Psychological Review*, *56*(1), 51-65. doi:[10.1037/h0062474](https://doi.org/10.1037/h0062474).
- Heckers, S., Zalesak, M., Weiss, A., Ditman, T., & Titone, D. (2004). Hippocampal Activation During Transitive Inference in Humans. *Hippocampus*, *14*(2), 153-162. doi:[10.1002/hipo.10189](https://doi.org/10.1002/hipo.10189).
- Herrnstein, R. J. and E. G. Boring (1965). *A Sourcebook in the History of Psychology*. Cambridge, MA, Harvard University Press.
- Holyoak, K., & Patterson, K. (1981). A positional discriminability model of linear-order judgments. *Journal of Experimental Psychology: Human Perception and Performance*, *7*(6), 1283-1302. doi:[10.1037/0096-1523.7.6.1283](https://doi.org/10.1037/0096-1523.7.6.1283).
- Hunter, W. (1920). The temporal maze and kinaesthetic sensory processes in the white rat. *Psychobiology*, *2*(1), 1-17. doi:[10.1037/h0073855](https://doi.org/10.1037/h0073855).
- Jou, J. (1997). Why is the alphabetically middle letter in a multiletter array so hard to determine? Memory processes in linear-order information processing. *Journal of Experimental Psychology: Human Perception and Performance*, *23*(6), 1743-1763. doi:[10.1037/0096-1523.23.6.1743](https://doi.org/10.1037/0096-1523.23.6.1743).
- Keppel, G., L. Postman, et al. (1968). Studies of learning to learn: VIII. The influence of massive amounts of training upon the learning and retention of paired-associate lists. *Journal of Verbal Learning and Verbal Behavior*, *7*(4), 790-796. doi:[10.1016/S0022-5371\(68\)80143-4](https://doi.org/10.1016/S0022-5371(68)80143-4).
- Lashley, K. S. (1951). The problem of serial order in behavior. *Cerebral mechanisms in behavior*. L. A. Jeffries. New York, NY, John Wiley & Sons: 112-136.
- Marshuetz, C., P. A. Reuter-Lorenz, et al. (2006). Working memory for order and the parietal cortex: an event-related fMRI study. *Neuroscience*, *139*, 311-316. doi:[10.1016/j.neuroscience.2005.04.071](https://doi.org/10.1016/j.neuroscience.2005.04.071).
- McGonigle, B., & Chalmers, M. (1977). Are monkeys logical? *Nature*, *267*(5613), 694-696. doi:[10.1038/267694a0](https://doi.org/10.1038/267694a0).
- Meck, W., & Church, R. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behavior Processes*, *9*(3), 320-334. doi:[10.1037/0097-7403.9.3.320](https://doi.org/10.1037/0097-7403.9.3.320).
- Miller, R. R. (2006). Challenges facing contemporary associative approaches to acquired behavior. *Comparative Cognition & Behavior Reviews*, *1*, 77-93. doi: [10.3819/cabr.2008.10005](https://doi.org/10.3819/cabr.2008.10005)
- Moyer, R., & Landauer, T. (1967). Time required for judgments of numerical inequality. *Nature*, *215*(5109), 1519-1520. doi:[10.1038/2151519a0](https://doi.org/10.1038/2151519a0).
- Nieder, A. (2005). Counting on neurons: The neurobiology of numerical competence. *Nature Reviews Neuroscience*, *6*(3), 177-190.
- Nieder, A. and E. K. Miller (2004). A parieto-frontal network for visual numerical information in the monkey. *Proceedings of the National Academy of Sciences*, *101*(19), 7457-7462.
- Olton, D., & Samuelson, R. (1976). Remembrance of places passed: Spatial memory in rats. *Journal of Experimental Psychology: Animal Behavior Processes*, *2*(2), 97-116. doi:[10.1037/0097-7403.2.2.97](https://doi.org/10.1037/0097-7403.2.2.97).
- Pavlov, I. P. (1927). *Conditioned Reflexes*. London, UK, Oxford University Press.
- Premack, D. (1976). *Intelligence in Ape and Man*. Hillsdale, NJ, Lawrence Erlbaum Associates.
- Roberts, W. (1998). *Principles of Animal Cognition*. Boston, The McGraw-Hill Companies.
- Rumbaugh, D. M. (1977). *Language learning by a chimpanzee: The Lana project*. New York, NY, Academic Press.
- Skinner, B. F. (1938). *The behavior of organisms*. New York, NY, Appleton-Century-Crofts.
- Sternberg, S. (1975). Memory scanning: New findings and current controversies. *The Quarterly Journal of Experimental Psychology*, *27*(1), 1-32.
- Straub, R., Seidenberg, M., Bever, T., & Terrace, H. (1979). Serial learning in the pigeon. *Journal of the Experimental Analysis of Behavior*, *32*(2), 137-148. doi:[10.1901/jeab.1979.32-137](https://doi.org/10.1901/jeab.1979.32-137).
- Terrace, H. (1979). Linguistic Apes: What Are They Saying? New York University *Education Quarterly* *9*(1).
- Terrace, H. S. (1984). Animal Cognition. *Animal Cognition*. H. L. Roitblat, T. G. Bever and H. S. Terrace. Hillsdale, NJ, Lawrence Earlbaum Associates: 7-28.
- Terrace, H. S. (1984). Simultaneous chaining: The problem it poses for traditional chaining theory. *Quantitative Analyses of Behavior: Discrimination Processes*. M. L. Commons, R. J. Herrnstein and A. R. Wagner. Cambridge, MA, Ballinger Publishing Co.: 115-138.
- Terrace, H. S., L. Son, et al. (2003). Serial expertise of rhesus macaques. *Psychological Science*, *14*(1), 66-73.
- Tolman, E. (1948). Cognitive maps in rats and men. *Psychological Review*, *55*(4), 189-208. doi:[10.1037/h0061626](https://doi.org/10.1037/h0061626).
- Tolman, E., Ritchie, B., & Kalish, D. (1992). Studies in spatial learning: I. Orientation and the short-cut. *Journal of Experimental Psychology: General*, *121*(4), 429-434. doi:[10.1037/0096-3445.121.4.429](https://doi.org/10.1037/0096-3445.121.4.429).
- Treichler, F., Raghanti, M., & Van Tilburg, D. (2003). Linking of serially ordered lists by macaque monkeys (*Macaca mulatta*): List position influences. *Journal of Experimental Psychology: Animal Behavior Processes*, *29*(3), 211-221. doi:[10.1037/0097-7403.29.3.211](https://doi.org/10.1037/0097-7403.29.3.211).

- Zorzi, M., K. Priftis, et al. (2006). The spatial representation of numerical and non-numerical sequences: Evidence from neglect. *Neuropsychologia*, *44*(7), 1061-1067. doi: [10.1016/j.neuropsychologia.2005.10.025](https://doi.org/10.1016/j.neuropsychologia.2005.10.025).
- Yates, F. A. (1966). *The Art of Memory*. Chicago, IL, The University of Chicago Press.