The simultaneous chain: a new approach to serial learning

Herbert S. Terrace

Department of Psychology and Psychiatry, Columbia University, New York, NY 10027;
and the New York State Psychiatric Institute, USA

Recent advances have allowed the application of behaviorism's rigor to the control of complex cognitive tasks in animals. This article examines recent research on serially organized behavior in animals. 'Chaining theory', the traditional approach to the study of such behavior, reduces intelligent action to sequences of discrete stimulus-response units in which each overt response is evoked by a particular stimulus. However, such theories are too weak to explain many forms of serially organized cognition, both in humans and animals. By training non-human primates to produce arbitrary sequences that cannot be learned as chains of particular motor responses, the simultaneous chaining paradigm has overcome limitations of chaining theory in experiments on serial expertise, the use of numerical rules, knowledge of ordinal position, and distance and magnitude effects.

Introduction

The psychology of the conditioned response is arguably behaviorism's major achievement. Sophisticated theories of conditioning [1] have been applied to a wide range of behavior and those principles have led to the discovery of neural mechanisms of conditioned behavior at cortical [2], sub-cortical [3] and cellular [4] levels of the brain. Unfortunately for behaviorism, its major achievement exposed its major limitation. Intelligent behavior is greater than the sum of discrete conditioned responses. Although it was once hoped that chaining theory could explain how individually conditioned responses could be linked together to form complex sequences [5], there is ample theoretical and empirical evidence that many forms of serially organized behavior are beyond its grasp (see Box 1). What is needed is an animal model of serially organized behavior that is based on learning ordinal relationships between stimuli rather than on the conditioning of new responses.

One of the main attractions of behaviorism is its rigorous, non-verbal methodology. There is, however, nothing inherent in that methodology that restricts its application to research on conditioning. Recent advances in technology have facilitated the application of behaviorism's rigor to the control of complex cognitive tasks in animals [6]; for example, concept formation [7], timing [8],

the discrimination of numerical quantities [9] and short-term memory [10]. Here I will focus on serial memory, a topic that has been particularly refractory to study in animals. The reasons are many but two are of special relevance to this article. The first is the lack of a suitable method for measuring an animal's memory of more than a single item on a given trial. The second is the view that the ability to execute arbitrary sequences requires language.

Ironically, the main impetus for recent interest in serially organized behavior in animals is the now discredited claim that apes (mainly chimpanzees) could create grammatical sequences [11,12] and by the success of a new paradigm for training monkeys to produce complex arbitrary sequences [13] without the benefit of any 'linguistic' training. As I will make clear below, those sequences differ in many important respects from the kind of chained sequences of discrete stimulus–response (S–R)

Box 1. Chaining theory

Staring with Ebbinghaus [28], various forms of chaining theory have assumed that an understanding of learned sequences in animals and humans would follow directly from an understanding of how particular stimuli became associated with particular responses. On this view all instances of serially organized behavior are reducible to discrete S–R units, each unit linked to the next by virtue of extensive practice. It matters not whether the sequence in question is a sentence or tying a shoelace [50–52].

In one of the most influential articles in modern psychology, Lashley [53] observed that chaining theory could not account for knowledge of relationships between non-adjacent items in serially organized behavior and presented two important examples to support his argument. First, he noted that all human languages assume knowledge of relationships between words from different parts of a sentence. For example, if someone told you that the girl wearing the blue uniform won the race, you would understand that it was the girl, and not the 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 are 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 has had less influence on animal cognition than it has had 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 arbitrary than human language [6]. That view is no longer tenable. Recent advances in our understanding of serially organized behavior in animals [54,55], have confirmed that Lashley's criticisms of chaining theory apply with the same force to animal behavior as it does to human behavior.

Corresponding author: Terrace, H.S. (terrace@columbia.edu).

Box 2. Successive versus simultaneous chains

According to chaining theory, what a subject learns in a maze with n choice points, are n discrete stimulus–response (S–R) associations. Consider the maze in Figure I in which $n=7$. In this instance, the subject only needs to learn 7 associations to get to the goal: $[S_1:R_1], [S_2:R_2]...[S_7:R_7]$. Knowledge of those associations would allow the subject to execute the following successive chain:
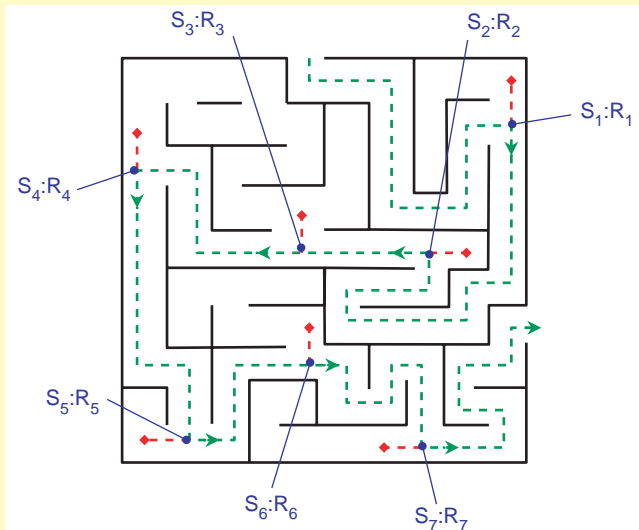

Figure I. (see text)

$S_1:R_1 \rightarrow S_2:R_2 \rightarrow S_3...S_7:R_7 \rightarrow S^R$ [56]. Crucially, it also insures that, having responded correctly to S_n, that stimulus disappears and that the subject will encounter S_{n+1}, and only S_{n+1}, at the next choice point. Thus, when the subject is at S_3, cues from the choice points S_1, S_2, S_4, S_5, S_6 and S_7 cannot compete for the subject's attention.

Simultaneous chaining paradigm (SCP)

The sine qua non of a simultaneous chain is that all list items are displayed throughout each trial and that no *differential* feedback is provided following each response as to the identity of the next item. All correct responses provide uniform feedback in that they allow a trial to continue. Similarly, 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.

Consider, for example, a trial on which a subject is required to produce a 7-item simultaneous chain [14]. The 7 stimuli are presented simultaneously (typically, on a touch-sensitive video monitor) until the subject either makes an error or earns a reward by responding to those stimuli in the correct order. A second crucial difference between a successive and a simultaneous chain is the spatial location of the choice points on each trial. On a successive chain, the spatial location of choice points remains fixed throughout training. That allows a subject to learn the maze as a specific sequence of motor responses. On a simultaneous chain, the spatial configuration of choice points is changes randomly from trial to trial. As a consequence, subjects cannot learn the required sequence as a fixed set of motor responses. The example in Figure II illustrates the differences in the paths a subject would have to follow to respond correctly on three sample trials of a 7-item simultaneous chain.

$$A \rightarrow B \rightarrow C \rightarrow D \rightarrow E \rightarrow F \rightarrow G$$

Figure II. (see text)

units on which animals have been trained for more than 100 years, such as running through a maze. I will refer to sequences composed of discrete S–R units as 'successive' chains because opportunities for making correct responses (and errors) are spatially and temporally isolated from one another [14]. I will refer to sequences trained by the new paradigm as 'simultaneous' chains because all of the choice points are presented simultaneously and it is

therefore possible to skip forward or backward to make errors to any of the other choice points throughout the entire trial. These and other basic differences between successive and simultaneous chains are elaborated in Box 2.

Experiments on simultaneous chains with non-human primates (the simultaneous chaining paradigm, or SCP) have provided opportunities for investigating a wide range
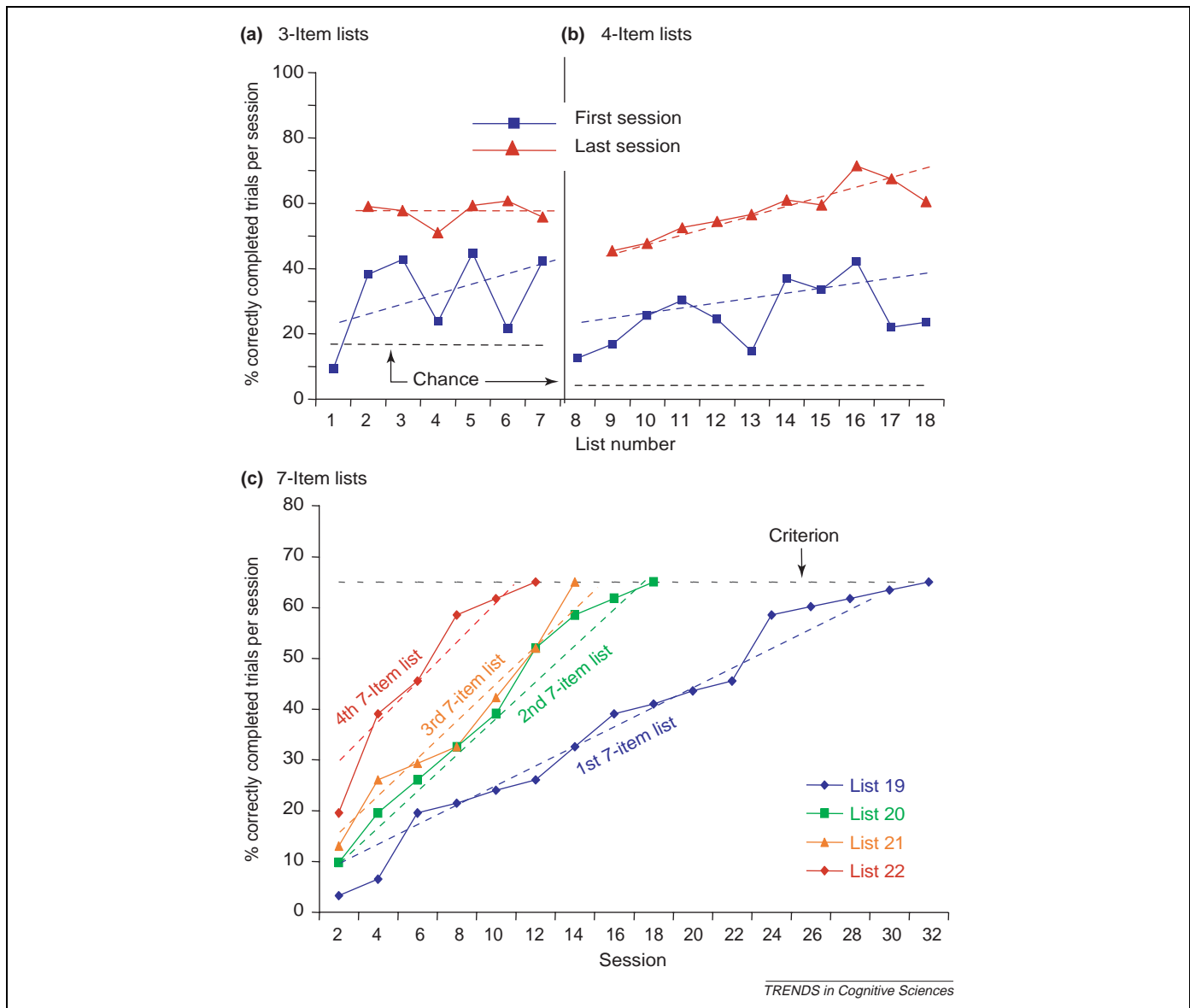
Figure 1. Learning curves showing the development of serial expertise in monkeys [15]. **(a,b)** The percentage of correctly completed trials during the first and the last sessions of training on **(a)** 3-item lists and **(b)** 4-item lists. The 3-item lists were the first on which the monkeys were trained, and they were given a maximum of three sessions of training on any one list. The slopes of the regression lines shown for first-session accuracy during successive 3- and 4-item lists differed significantly from zero. **(c)** The mean accuracy of responding on each 7-item list during even-numbered sessions. Each list was trained to an accuracy criterion of 65% correctly completed trials during one session (hence the abscissa here is Session, not List number). The probability of guessing the correct sequence on the first trial of training was calculated using the conservative assumption that there were no backward errors (an assumption that is empirically true for monkeys trained on multiple lists). On new 3-, 4- and 7-item lists those probabilities were, respectively, 0.17, 0.04 and 0.0005.

of human cognitive phenomena that are beyond the scope of traditional chaining theory. These include serial expertise [15], knowledge of the ordinal position of list items [15–17], numerical sequences [18–20] and distance and magnitude effects [21–23]. This article will summarize some recent research on those topics and indicate how it has led to animal models of cognitive phenomena that, until recently, have been regarded as exclusively human.

Serial expertise
In his classic experiment on learning sets, Harlow [24] showed that it was easier for a subject to induce a general strategy for mastering a particular cognitive task ('win-stay; lose-shift') by working on many exemplars of a particular problem than by waiting for a subject to

achieve a high level of accuracy on a specific exemplar. Harlow's strategy was applied in an experiment in which 4 monkeys were trained to produce multiple 3-, 4- and 7-item simultaneous chains (referred to hereafter as 'lists') [15]. Each list consisted of photographs of natural objects or events that were presented on a touch-sensitive video monitor. The monkeys, none of whom had any previous list-learning experience, were trained on each new 3- and 4-item list for a maximum of 3 sessions. The idea was to induce a strategy for identifying the ordinal position of items on new lists by varying the problem and by gradually increasing its difficulty.

As shown in Figure 1, the learning-set strategy worked extremely well. On 3-item lists, the percentage of correctly completed trials increased progressively during the first
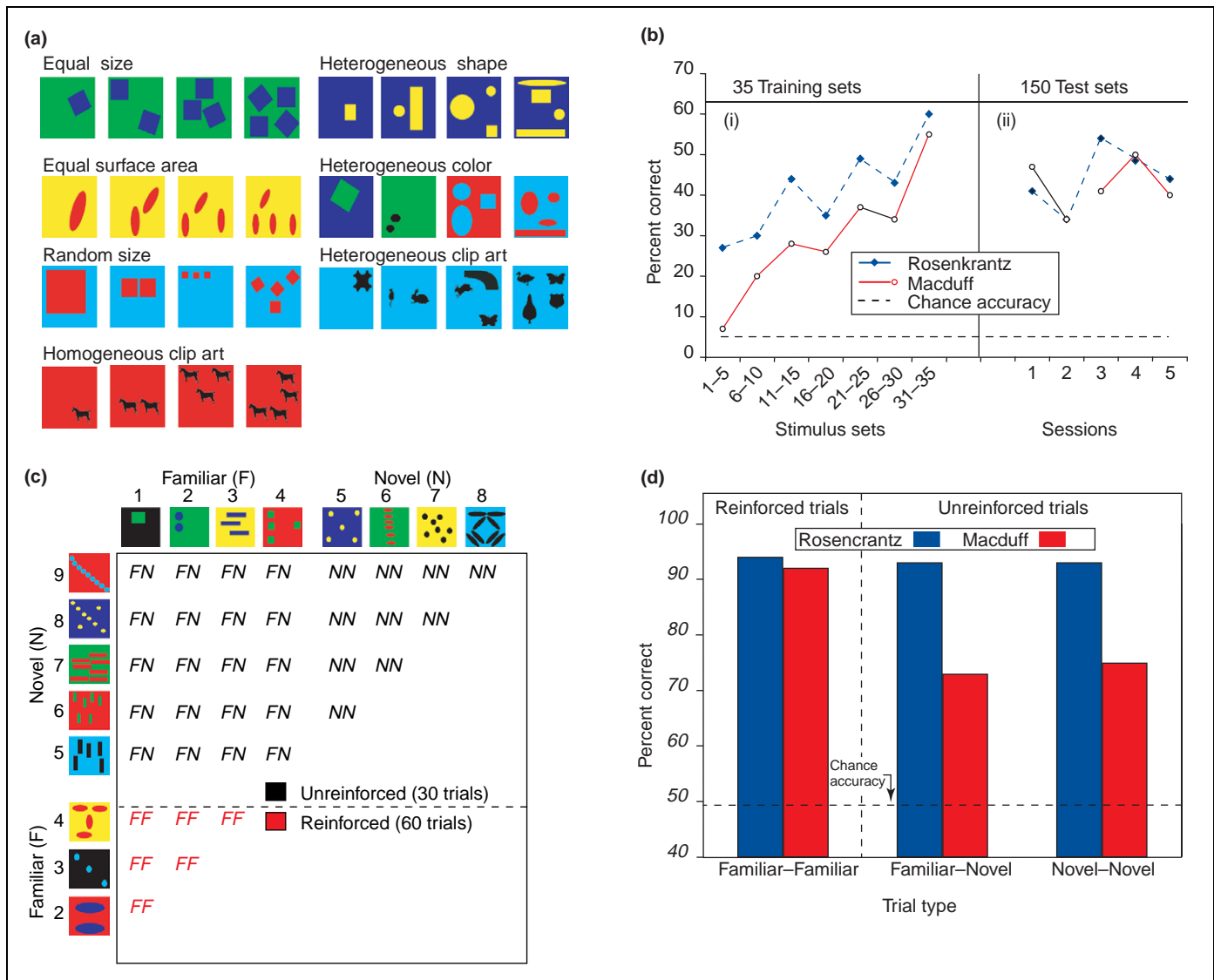
Figure 2. Assessing the numerical abilities of rhesus monkeys [18]. **(a)** Exemplars of the seven different types of stimulus sets used to train monkeys to produce ascending numerical lists. Each set controlled for a particular non-numerical dimension that could have been used as the basis for discriminating stimuli containing a different number of elements (e.g. size, shape, color, area, and homogeneity of elements). **(b)** Accuracy of the two monkeys, Rosenkrantz and Macduff, on numerical lists during training and testing. Left panel: percentage of correctly completed trials during the first session for each of 35 training sets in blocks of five sessions. Right panel: percentage of correctly completed trials on 150 trial-unique novel test sets. **(c)** The 36 pairs of the numerosities 1–9 used to test a monkey's on their knowledge of numerical order following training on the ascending sequence 1→2→3→4. The numerosities 5–9 were presented for the first time during this test. Each numerical pair was defined with respect to the monkeys' previous experience with the constituent numerosities: Familiar–Familiar (FF), Familiar–Novel (FN), Novel–Novel (NN). **(d)** Accuracy of responding to numerical subsets drawn from each of the three subset categories shown in (c).

session of training on each new list. This being a new task on which the monkeys had no previous list training, the level of accuracy at the end of training on each list reach a plateau of ∼50% correctly completed trials. On 4-item lists, the percentage of correctly completed trials increased progressively during the first and the last sessions of training on each list. In the case of 7-item lists, which were trained to a criterion of completing at least 65% of the trials in a single session, monkeys needed progressively fewer sessions to learn successive lists. Indeed, on their fourth 7-item list, the monkeys identified the first two items with the fewest possible logical guesses [15].

The lists trained in this study are arguably the most difficult serial tasks mastered by a non-human primate. Those include the production of symbolic, [11,12,25], numerical [19,25] and arbitrary lists [15] and the

recognition of the ordinal position of items on arbitrary lists [17,26]. As I have argued elsewhere [27], the performance of the monkeys that learned 7-item lists compares favorably with that of a person attempting to deduce, by trial and error, a 7-digit PIN to operate an ATM on which the position of each number varies randomly from trial to trial. It is doubtful, however, that the sequential skills of these monkeys reflect the upper limit of their serial expertise. The steady decrease in the number of sessions needed to master new lists suggests that they could learn such lists more rapidly and that they could also master longer lists.

Numerical ability: rule-governed simultaneous chains
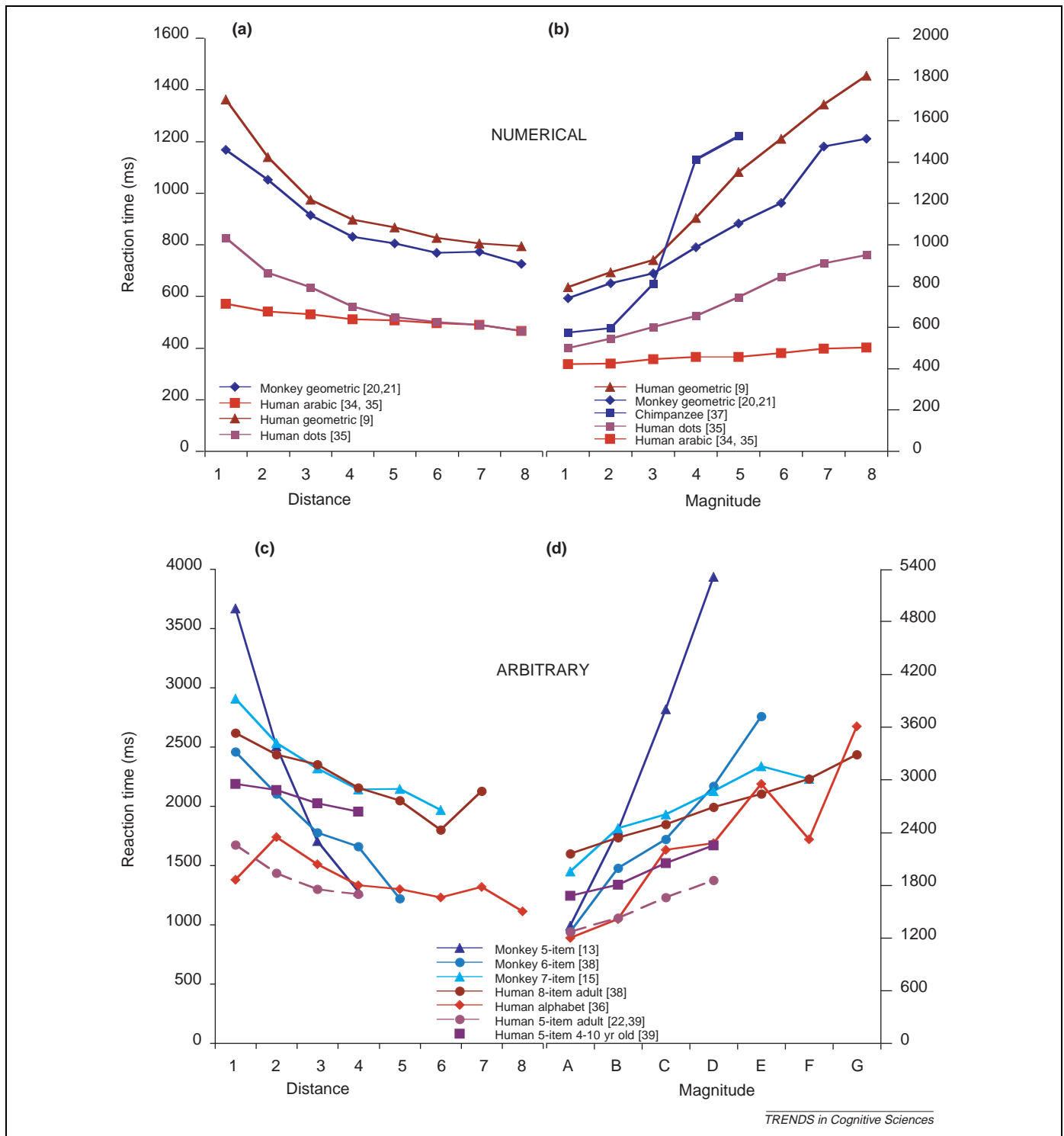Trial-and-error learning is necessary to master lists composed of arbitrary items. There is, however, no reason

Figure 3. Distance and magnitude functions obtained from human subjects and monkeys. These functions were obtained on subtest tests composed of **(a,b)** numerical and **(c,d)** arbitrary items. With the exception of the experiment on numerical judgments by a chimpanzee [37], all of the non-human primates were first trained using the SCP to learn lists of arbitrary or numerical stimuli. 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 and letters of the alphabet or with arbitrary or numerical stimuli similar to those used in the experiments with non-human primates.

why list items must be arbitrary. They could just as well conform to a particular rule. Consider, for example, a simultaneous chain on which the subject is required to respond, in an ascending order, to different sets of geometric stimuli, each containing a different number of elements of different shapes, sizes and colors. Such a list could be learned by trial and error *or* by applying the

ascending rule. If it were learned by trial and error, new lists composed in the same manner should be equally difficult. By contrast, subjects that learned the ascending rule could execute new lists correctly on the first trial on which they were presented.

Precisely such results were obtained in an experiment in which monkeys were trained to respond to

geometrically defined exemplars of the numerosities 1, 2, 3 and 4 in an ascending order [18]. Examples of the numerical exemplars used in this experiment are shown in Figure 2a. After training on 35 four-item lists, each composed of a different set of exemplars, monkeys were tested on 150 novel lists, all trial unique. As shown in Figure 2b, the monkeys learned to respond at a high level of accuracy on the 35 training lists and continued to do so, without any decrement, on the 150 novel lists. As there is no benefit of memorizing the order of list items that were trial unique, it follows that the monkeys abstracted a numerical rule during their initial training on numerically defined lists.

By itself, this result does not support the conclusion that monkeys can learn an ordinal rule. Stronger evidence of a monkey's ability to use an ordinal rule was obtained on a test composed of all of the 36 numerical pairs that can be derived from the numerosities 1–9. These are shown in Figure 2c. Ten of those pairs were composed of novel numerosities whose values ranged from 5–9. Both monkeys responded correctly on more than 70% of the trials on which both items were novel (cf. Figure 2b). As neither subject in this experiment had any previous experience with the numerosities 5–9, their ability to respond appropriately to those numerosities shows clearly that a monkey can extrapolate an ordinal rule to determine the correct order in which to respond to novel numerosities.

Knowledge of ordinal position

When a monkey learns an arbitrary list, does it learn item-order information (associations between an item and its ordinal position) in addition to item–item information (associations between successive list items)? That issue has been investigated in a variety of experiments in which monkeys, trained on different arbitrary lists, were given problems that would be difficult, if not impossible, to solve using item–item information.

Derived lists

In an experiment based on Ebbinghaus' method of derived lists [28], two monkeys were first trained to produce four lists, each composed of novel photographs ($A_1 \rightarrow B_1 \rightarrow C_1 \rightarrow D_1$, $A_2 \rightarrow B_2 \rightarrow C_2 \rightarrow D_2$, $A_3 \rightarrow B_3 \rightarrow C_3 \rightarrow D_3$ and $A_4 \rightarrow B_4 \rightarrow C_4 \rightarrow D_4$; subscripts refer to list number) [16]. They were then trained on four new 4-item lists that were derived from the original lists, with the constraint that only one item from any of the four original lists could be used in any of the new lists. Two of the derived lists maintained each item's original ordinal position ($A_2 \rightarrow B_4 \rightarrow C_1 \rightarrow D_3$ and $A_3 \rightarrow B_1 \rightarrow C_4 \rightarrow D_2$). On the other two lists, the ordinal positions of all items were changed ($B_3 \rightarrow A_1 \rightarrow D_4 \rightarrow C_2$ and $D_1 \rightarrow C_2 \rightarrow B_3 \rightarrow A_4$).

Each derived list should be equally difficult to learn if a monkey's knowledge of the four original lists were limited to item–item associations. If, however, a monkey acquired knowledge of each item's ordinal position while learning the original lists, the two maintained lists should be easier to learn than the two changed lists. The latter prediction was confirmed. Both monkeys acquired the two maintained lists with virtually no errors, but it took them as long to acquire either of the changed lists as it would to

learn a list composed of novel items. Similar results have been obtained in an experiment in which human subjects were trained with the derived-list paradigm [28,29].

Errors based on ordinal category

A monkey's knowledge of ordinal position has also been evaluated by determining the confusability of list items that occupy the same ordinal position on different lists [17]. In one experiment, monkeys were first trained on 10 different lists, each composed of three different fractal images. During the study portion of each trial, items from a given list were displayed successively. During the test, the list items were displayed simultaneously (as in the SCP), together with a distractor item that was chosen at random from the remaining 27 fractal images. The monkey's task was to respond to the items presented during the study portion of the trial in the order in which they were presented.

If monkeys only formed item–item associations between the fractal images from a given list, errors to the distractor should be equally probable with respect to the subject's first, second and third responses. Contrary to that hypothesis, there was a high frequency of errors to distractors that occupied the same ordinal position as the item from the triplet that was displayed on a given trial. For example, if the sample sequence was $D_1 \rightarrow D_2 \rightarrow D_3$, and if the distractor was G_3, the subject would make many more errors at the third position than at the first or second positions. In agreement with the experiment on derived lists [16], this suggests that the monkey's natural strategy was to encode list items with respect to their ordinal position on each list.

Distance and magnitude effects

When comparing two items on a particular dimension, for example, size, the more discrepant the items, the more accurate and the more rapid the comparison. Such distance effects were originally observed in experiments on psychophysics and human perception, such as in judgments of line length [30], but they have also been observed in experiments on human memory [31]. The latter gave rise to what has come to be known as the 'symbolic distance effect' (SDE), in which distance is defined psychologically. The first demonstration of the SDE occurred in a deceptively simple experiment [32]. When adult human subjects were asked to select the larger of a pair of Arabic numbers, their reaction times (RTs) decreased progressively as the distance between the numbers increased. Similar results were obtained when human subjects were asked to judge the relative magnitude of displays containing random arrays of dots [9,33]. SDEs have also been obtained with stimuli from arbitrary continua, for example, letters of the alphabet [34], and from experiments in which subjects were asked to compare two mental images to decide which one was larger [35].

In addition to distance, judgments of relative magnitude are affected by the magnitude of the smaller item. When the distance between the smaller and the larger item is held constant, RTs of judgments of relative magnitude increase as the magnitude of the smaller

item increases. Thus, when human subjects are asked which number is larger, 7 or 8, RTs are longer than when the same question is asked about the numbers 3 and 4 [33]. The relationship between the smallest difference between two stimuli that can be discriminated and the smaller of the two stimuli, famously known as Weber's law in the case of natural continua ($\Delta I/I = k$), has been referred to as a 'magnitude effect' in the case of psychological stimuli.

Distance and magnitude effects have been obtained from non-human primates in experiments in which monkeys were tested on their ability to order 2-item subsets drawn from items that comprised arbitrary or numerical simultaneous chains on which they were previously trained. The results of these experiments, which are strikingly uniform, are shown in Figures 3a–d. Figures 3a and b show distance and magnitude effects obtained with numerical stimuli; Figures 3c and d, with arbitrary stimuli.

The function labeled 'Monkey geometric' in Figures 3a and b is based on the median RTs that were obtained in the previously described experiment with rhesus macaques on numerical discrimination [21], and from a replication of that experiment in which a squirrel monkey *(Saimiri sciareus)* and a baboon *(Papio hamadrayas)* [20] were trained with the same numerical stimuli (cf. Figure 2a). The function labeled 'Human geometric' was obtained from an experiment with college students who were also tested with the same numerical stimuli [9]. Also shown in Figure 3b is a magnitude function obtained from Ai, a chimpanzee trained to report the number of dots displayed on a computer monitor by selecting the appropriate Arabic number on a computer keyboard [36]. Taken together these data indicate that reliable distance and magnitude functions can be obtained from non-human primates that are trained and tested with numerical stimuli.

Similarly congruent distance and magnitude functions were obtained from experiments in which non-human primates and adult and juvenile humans were trained using the SCP to learn arbitrary sequences composed of photographs or geometric forms [22,23,37–39]. These are shown in Figures 3c and d. A comparison of Figures 3a and b with 3c and d reveals an intriguing difference in the absolute values of the RTs obtained with arbitrary and numerical stimuli. RTs to arbitrary stimuli (Figures 3c and d) were roughly twice as long as RTs to numerical stimuli (Figures 3a and b). Assuming that comparisons of both types of stimuli are made in working memory, it appears that the longer RTs for arbitrary stimuli include the time needed to retrieve information about the ordinal position of such stimuli from long-term memory. That does not seem to be the case for numerical stimuli whose ordinal position can be determined algorithmically in working memory [40].

Distance effects obtained with arbitrary stimuli have two important theoretical implications. Chaining theory would predict an *increase* in RT with increasing distance between arbitrary stimuli because associative strength decreases as the ordinal distance between items increases. As can be seen in Figures 3a and c, the slopes of the distance function are the exact opposite of what chaining theory would predict. Of broader interest are the implications of the data shown in Figures 3a–d for theories of representation. Many computational theories claim that cognitive representations are propositional, that is, they are based upon rules of syntax and semantics [41,42]. In experiments with human subjects, it is difficult to tease apart the influences of propositional thinking, which requires language, from the influence of perceptual–analogical processing, which may not [43]. A monkey's knowledge of the ordinal position of arbitrary items cannot, of course, be based on linguistic representations. A more plausible candidate is analog coding of ordinal position, which can account for knowledge of ordinal position and distance effects in the case of arbitrary stimuli [43]. Analog coding could also account for RTs of judgments of the numerical magnitude of stimuli. Interestingly, recent experiments on the neural basis of numerical judgments by monkeys suggest a continuum of analog cognitive–perceptual processes that are shared by human and non-human primates (see Box 3).

Conclusion

The simultaneous chaining paradigm provides an important bridge for comparing animal and human cognition. Its success in training animals to order arbitrary and numerical stimuli stems from the absence of any requirement for subjects to learn new motor responses when learning a new sequence. In addition to the experiments described in this article, the SCP has provided a basis for performing experiments on cognitive imitation [44], on metacognition [45], and on neural mechanisms that mediate serially organized behavior [26,46–49]. Given

---

Box 3. Neural mechanisms of serially organized behavior

Paradigms for training serial skills in animals have led to recent advances in our understanding of their neural structure and organization, in particular, the neural basis of numerical discrimination and the temporal planning of arbitrary sequences. Using abstract numerical stimuli (similar to those in Figure 2a in main text), Nieder and Miller [49,57] trained monkeys to discriminate numerosities whose value ranged from 1 to 5. They then showed that specific cells in the pre-frontal cortex responded maximally to particular numerical values and that the bandwidth of each type of cell increased progressively as the value of the numerosity to which it was 'tuned' increased. For example, cells that responded maximally to stimuli containing two geometric elements responded to a narrower range of other numerical values than a cell that responded maximally to stimuli containing five geometric stimuli.

That property of numerically specific cells might explain why numerical discrimination improves with increasing numerical distance (an example of a 'symbolic distance effect') and why, at a given numerical distance, the discriminability of two numerical values decreases as numerical size increases (an example of Weber's Law). An experiment by Ninokura *et al.* [48] identified cells in the lateral prefrontal cortex (LPC) of monkeys that encoded information needed to plan a 3-item simultaneous chain. The monkeys in this experiment had to remember the temporal order of three successively presented visual stimuli (A, B and C) that were presented successively, and then respond to them in the same order when they were presented simultaneously. During a 1.5 s interval that preceded the simultaneous test, cells in LPC responded differentially to each of the six possible sequences during the study portion of that trial (ABC, ACB, BAC, BCA, CAB and CBA). This is strong evidence that, at least for 3-item sequences, a subject plans the entire sequence before executing it.

Box 4. Future directions

Planning

Representations, not matter how complex, are of no use unless they are translated into action. This requires serial planning of motor movements. Variations of the SCP that rearrange the physical configuration of list items should help to determine how far ahead a subject plans when executing a sequence [37].

Chunking

Can monkeys combine small (3-item) lists they have already mastered into larger lists by using each 3-item list as chunk? A monkey's ability to use chunks can also be studied in experiments in which adjacent items on a particular list are photographs of objects from the same category (e.g. fruit, birds trees, etc.) [58].

Transitive inference

The transitive inference (TI) paradigm [59–62] has generated an extensive literature on ordinal knowledge of human children and non-human primates. With the TI paradigm, a subject is first trained on individual ordinal relationships, for example, A>B, B>C, C>D, D>E... and is then tested on their knowledge of the ordinal relationships between non-adjacent pairs, for example, B and D. It is of interest to compare ordinal representations of arbitrary stimuli established by the SCP and the TI paradigms. Similar outcomes would suggest a natural 'order' line of the type suggested for number [63].

Memory of multiple list items

To date, all experiments on short-term serial memory in monkeys test only for memory of a single item following the presentation of a multi-item list [10]. The SCP provides a means of testing a subject's memory of all of the items presented during the sample by presenting all of those items simultaneously, along with various distractors. Memory of item information can be tested by requiring subjects to respond to all of the sample items, in any order; memory of order information, by requiring subjects to respond to the sample items in the order in which they were presented. Such experiments would provide a valuable basis for comparing short-term serial memory in human and non-human primates.

---

that many of the discoveries about simultaneous chaining have been serendipitous, I anticipate that the future of research on this topic (see also Box 4) will contain various surprises that fill many gaps in our understanding of how the verbal human mind evolved from the poorly understood non-verbal animal mind.

References

1 Miller, R.J. and Escobar, M. (2002) Laws and models of basic conditioning. In *Stevens' Handbook of Experimental Psychology* (Gallistel, C.R., ed.), pp. 47–102, John Wiley & Sons

2 Bao, S. *et al.* (2005) Cerebellar cortical inhibition and classical eyeblink conditioning. *Proc. Natl. Acad. Sci. U. S. A.* 99, 1502–1507

3 Lanuza, E. *et al.* (2004) Unconditioned stimulus pathways to the amygdala: effects of posterior thalamic and cortical lesions on fear conditioning. *Neuroscience* 125, 305–315

4 Antonov, I. *et al.* (2001) The contribution of activity-dependent synaptic plasticity to classical conditioning in Aplysia. *J. Neurosci.* 21, 6413–6422

5 Dayan, P. (2002) Reinforcement learning. In *Stevens' Handbook of Experimental Psychology (Vol. 3: Learning, Motivation & Emotion)* (Gallistel, C.R., ed.), pp. 103–129, John Wiley & Sons

6 McGonigle, B. and Chalmers, M. (2002) A behavior based fractionation of cognitive comptence with applications: A comparative approach. *Int. J. Comp. Psychol.* 15, 75–82

7 Wasserman, E. *et al.* (2001) Same–different conceptualization by baboons. *J. Comp. Psychol.* 115, 42–52

8 Church, R. (2002) Temporal learning. In *Stevens' Handbook of Experimental Psychology (Vol. 3: Learning, Motivation and Emotion)* (Gallistel, C.R., ed.), pp. 365–393, John Wiley & Sons

9 Brannon, E. and Terrace, H. (2002) The evolution and ontogeny of ordinal behavior. In *The Cognitive Animal* (Bekoff, M. *et al.*, eds), pp. 197–204, Oxford University Press

10 Wright, A.A. (2002) Monkey auditory list memory: Tests with mixed and blocked retention delays. *Anim. Learn. Behav.* 30, 158–164

11 Premack, D. (1976) *Intelligence in Ape and Man*, Erlbaum

12 Rumbaugh, D.M. (1977) *Language Learning by a Chimpanzee: The Lana Project*, Academic Press

13 D'Amato, M.R. and Colombo, M. (1988) Representation of serial order in monkeys (*Cebus apella*). *J. Exp. Psychol. Anim. Behav. Process.* 14, 131–139

14 Terrace, H.S. (1984) Simultaneous chaining: The problem it poses for traditional chaining theory. In *Quantitative Analyses of Behavior: Discrimination Processes* (Commons, M.L. *et al.*, eds), pp. 115–138, Ballinger Publishing Co

15 Terrace, H.S. *et al.* (2003) Serial expertise of rhesus macaques. *Psychol. Sci.* 14, 66–73

16 Chen, S. *et al.* (1997) Knowledge of the ordinal position of list items in rhesus monkeys. *Psychol. Sci.* 8, 80–86

17 Orlov, T. *et al.* (2000) Macaque monkeys categorize images by their ordinal number. *Nature* 404, 77–80

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

19 Matsuzawa, T. and Kawai, N. (2000) Numerical memory span in a chimpanzee. *Nature* 403, 39–40

20 Smith, B.R. *et al.* (2003) Numerity of a socially housed hamadryas baboon (*Papio hamadrayas*) and a socially housed squirrel monkey (*Saimiri sciareus*). *J. Comp. Physiol. Psychol.* 117, 217–225

21 Brannon, E.M. and Terrace, H.S. (2000) Representation of the numerosities 1-9 by rhesus monkeys. *J. Exp. Psychol. Anim. Behav. Process.* 26, 31–49

22 Colombo, M. and Frost, N. (2001) Representation of serial order in humans: A comparison to the findings with monkeys (*Cebus appela*). *Psychon. Bull. Rev.* 8, 262–269

23 Ohshiba, N. (1997) Memorization of serial items by Japanese monkeys, a chimpanzee, and humans. *Jpn. Psychol. Res.* 39, 236–252

24 Harlow, H.F. (1949) The formation of learning sets. *Psychol. Rev.* 56, 51–65

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

26 Carpenter, A.F. *et al.* (1999) Motor control encoding of serial position in a context-recall task. *Science* 283, 1752–1757

27 Terrace, H.S. (2002) Serial expertise and the evolution of language. In *The Transition to Language* (Wray, J.H. and Newmeyer, F.J., eds), pp. 64–90, Oxford University Press

28 Ebbinghaus, H. (1964) *Memory: A Contribution to Experimental Psychology*, Dover

29 Ebenholtz, S.M. (1963) Serial learning: position learning and sequential associations. *J. Exp. Psychol.* 66, 353–362

30 Welford, A.T. (1960) The measurement of sensory-motor performance: survey and reappraisal of twelve years progress. *Ergonomics* 3, 189–230

31 Potts, G.R. *et al.* (1978) Encoding and retrieval in comparative judgments. In *Cognitive Theory* (Vol. 3) (Castellan, N.J. and Restle, F., eds), pp. 243–308, Erlbaum

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

33 Buckley, P.B. and Gillman, C.B. (1974) Comparisons of digit and dot patterns. *J. Exp. Psychol.* 103, 1131–1136

34 Hamilton, J.M.E. and Sanford, A.J. (1978) The symbolic distance effect for alphabetic order judgements: A subjective report and reaction time analysis. *Q. J. Exp. Psychol.* 30, 33–43

35 Moyer, R.S. (1973) Comparing objects in memory: Evidence suggesting an internal psychophysics. *Percept. Psychophys.* 13, 180–184

36 Murofushi, K. (1997) Numerical matching behavior by a chimpanzee (*Pan troglodytes*): Subitizing and analog magnitude estimation. *Jpn. Psychol. Res.* 39, 140–153

37 Terrace, H. (2001) Comparative psychology of chunking. In *Animal Cognition and Sequential Behavior* (Fountain, S., ed.), pp. 23–56, Kluwer Academic Publishing

38 Guyla, M. and Colombo, M. (2004) The ontogeny of serial-order behavior in human (*Homo sapiens*) representation of a list. *J. Comp. Psychol.* 118, 71–81

39 Chalmers, M. and McGonigle, B. (1984) Are children any more logical than monkeys on the five term series problem? *J. Exp. Child Psychol.* 37, 355–377

40 Carey, S. (2004) Bootstrapping and the origin of concepts. *Daedalus* Winter, 59–68

41 Fodor, J.A. and Pylyshyn, Z.W. (1988) Connectionism and cognitive architecture: A critical analysis. In *Connections and Symbols* (Pinker, S. and Mehler, J., eds), pp. 3–72, MIT Press

42 Pylyshyn, Z. (1981) The imagery debate: Analogue versus tacit knowledge. *Psychol. Rev.* 88, 16–45

43 Kosslyn, S.M. (1994) *Image and Brain: The Resolution of the Imagery Debate*, MIT Press

44 Subiaul, F. *et al.* (2004) Cognitive imitation in rhesus macaques. *Science* 305, 407–410

45 Son, L. and Kornell, N. (2005) Metaconfidence judgments in Rhesus macaques: Explicit versus implicit mechanisms. In *The Missing Link in Cognition: Origins of Self-Reflective Consciousness* (Terrace, H.S. and Metcalfe, J., eds), Oxford University Press (in press)

46 Nieder, A. *et al.* (2002) Representation of the quantity of visual items in the primate prefrontal cortex. *Science* 297, 1708–1711

47 Grafton, S.T. *et al.* (1995) Functional mapping of sequence learning in normal humans. *J. Cogn. Neurosci.* 7, 497–510

48 Ninokura, Y. *et al.* (2004) Integration of temporal order and object information in the monkey lateral prefrontal cortex. *J. Neurophysiol.* 91, 555–560

49 Marshuetz, C. *et al.* (2000) Order information in working memory: fMRI evidence for parietal and prefrontal mechanisms. *J. Cogn. Neurosci.* 12, 130–144

50 Hull, C.L. (1932) The goal gradient hypothesis and maze learning. *Psychol. Rev.* 39, 25–43

51 Skinner, B.F. (1934) The extinction of chained reflexes. *Proc. Natl. Acad. Sci. U. S. A.* 20, 234–237

52 Skinner, B.F. (1957) *Verbal Behavior*, Appleton-Century-Crofts

53 Lashley, K.S. (1951) The problem of serial order in behavior. In *Cerebral Mechanisms in Behavior* (Jeffries, L.A., ed.), pp. 112–136, John Wiley & Sons

54 Fountain, S. *et al.* (2002) The organization of sequential behavior. In *Animal Cognition and Sequential Behavior* (Fountain, S.B. *et al.*, eds), pp. 115–150, Kluwer Academic Publishers

55 Terrace, H.S. (1993) The phylogeny and ontogeny of serial memory: List learning by pigeons and monkeys. *Psychol. Sci.* 4, 162–169

56 Skinner, B.F. (1938) *The Behavior of Organisms*, Appleton-Century-Crofts

57 Nieder, A. and Miller, E.K. (2004) A parieto-frontal network for visual numerical information in the monkey. *Proc. Natl. Acad. Sci. U. S. A.* 101, 7457–7462

58 Terrace, H.S. (2001) The comparative psychology of serially organized behavior. In *Avian Visual Cognition* (Cook, R., ed.), Comparative Cognition Press (available online: www.pigeon.psy.tufts.edu/avc/)

59 McGonigle, B.O. and Chalmers, M. (1977) Are monkeys logical? *Nature* 267, 355–377

60 McGonigle, B. (1987) Non-verbal thinking by animals? *Nature* 325, 110–112

61 Treichler, F.R. *et al.* (2003) Linking of serially ordered lists by macaque monkeys (*Macaca mulatta*): List position influences. *J Exp. Psychol. Anim. Behav. Process.* 29, 211–221

62 Boysen, S.T. *et al.* (1993) Processing of ordinality and transitivity by chimpanzees (*Pan Troglodytes*). *J. Comp. Psychol.* 107, 208–215

63 Dehaene, S. (1997) *The Number Sense: How the Mind Creates Mathematics*, Oxford University Press