Is the mind inherently forward looking?
Comparing prediction and retrodiction

JASON JONES AND HAROLD PASHLER
University of California, San Diego, La Jolla, California

It has been suggested that prediction may be an organizing principle of the mind and/or the neocortex, with cognitive machinery specifically engineered to detect forward-looking temporal relationships, rather than merely associating temporally contiguous events. There is a remarkable absence of behavioral tests of this idea, however. To address this gap, we showed subjects sequences of shapes governed by stochastic Markov processes, and then asked them to choose which shape reliably came after a probe shape (prediction test) or before a probe shape (retrodiction test). Prediction was never superior to retrodiction, even when subjects were forewarned of a forward-directional test.

Mining the data of past experience to predict future events is surely one of the most essential functions of the human mind and brain. The better our ancestors could predict when a predator would strike or when a tree would bear fruit, the better their chances for survival. Predictions based on past experience drive our choices about everything from choosing an entrée at a familiar restaurant to investing our money in the stock market. Given the utility of prediction, it would seem to be a reasonable conjecture that the brain may be highly adapted to generate predictions, rather than merely associating experiences or events that occur in close temporal proximity. The present study will examine this conjecture empirically.

The idea that memory is specialized for detecting forward-directed temporal relationships has sometimes been considered too obvious to need any demonstration. In 1887, in an article entitled "Why Do We Remember Forwards and Not Backwards?" the philosopher F. H. Bradley wrote, "Life being a process of decay and of continual repair, and a struggle throughout against dangers, our thoughts, if we are to live, must mainly go the way of anticipation. This, when we attend to it, seems quite evident and a mere commonplace." (Bradley, 1887, p. 581).

More recently, a number of neuroscientists have echoed Bradley's thesis, and asserted that prediction (rather than merely connecting or associating experiences) is the primary function of the neocortex (Hawkins & Blakeslee, 2004; Llinás, 2001). In the area of neurocomputational theory, popular approaches such as temporal difference learning (Sutton, 1988) assume an inherently directional learning process; the broad applicability of these frameworks to brain function has been enthusiastically advocated in recent years (see, e.g., Montague, Hyman, & Cohen, 2004; Rao & Sejnowski, 2003).

What is striking, however, is the absence of any behavioral test of the proposition that memory is specifically engineered to detect predictive relationships, as opposed to merely associating events that are temporally contiguous. At first blush, several research traditions would seem to bear on this. Many studies have compared "forward recall" to "backward recall" in paired associate learning. In this procedure, subjects are given a list of word or nonsense syllable pairs to memorize. Typically, the members of each pair are presented visually and at the same time, separated by a dash. (We will denote the pairs with letters; e.g., A–B.) Forward recall is tested using A as the probe (A–?), and backward recall is tested using B as the probe (?–B).

The general consensus in the literature has been that once certain methodological demands are satisfied, performance in the two tests is equivalent (a result commonly termed "associative symmetry"; Asch & Ebenholtz, 1962; Kahana, 2002). However, the implications for the question of whether prediction is superior to retrodiction are far from clear, for several reasons. First, in the typical paired associate study, both items in a pair (A and B) were at some point presented simultaneously. For example, in Murdock (1962), the stimuli were printed lists of word pairs. In Kahana (2002), complete word pairs were displayed on a computer monitor for 2 or more seconds. Jantz and Underwood (1958) presented first a nonsense syllable (A) and then repeated it in conjunction with the associated adjective, separated by a dash (A–B). It is difficult to answer the question of whether the direction in time in which items are presented affects the observer's association between the two items when the two items are in fact presented at the same time. Second, in paired associate learning, the associations to be learned are made explicit to the subject. The A and B items are presented as pairs, with specific instructions to the subject to memorize them as pairs. Thus, the results may say little about what spontaneous learning might be triggered by experiencing sequences.

Studies of verbal memory beyond the realm of paired associates do provide results that could be seen as suggesting an inherent advantage for prediction in memory. When a third word is added to a paired associates task (creating a three-word list: A–B–C), accuracy and reaction time (RT) advantages for recall are found in forward-cued conditions (i.e., AB?, _B?, and A?_) (Kahana & Caplan, 2002). In free recall of word lists, subjects are more likely to produce consecutive items in the same order as in the original studied list than in the reverse order (Kahana, 1996). Analysis of response times for serial forward or backward recall in word lists showed faster overall response for forward recall and evidence that backward recall was accomplished through repeated cycles of covert forward recall (Thomas, Milner, & Haberlandt, 2003). In contrast to the paired associate work, these studies appear to argue for asymmetry in recall performance.

Moving to the animal learning tradition, it has often been observed that Pavlovian conditioning is more reliably elicited when the conditioned stimulus (CS) precedes the unconditioned stimulus (US), rather than vice versa (e.g., Chang, Stout, & Miller, 2004; Spooner & Kellogg, 1947). This might seem to argue for a temporal asymmetry in the detection of temporal relationships. However, it might instead reflect a temporal asymmetry in the way that cognitive appraisals elicit emotions: For example, it would seem to be natural for an organism to fear a dreaded event that lies in its future, but not a dreaded event that lies in its past, given an equally strong belief in both.

What is needed, then, is to provide a proper test of whether predictive temporal relationships are preferentially detected. First, subjects must be presented with a controlled set of experiences, unfolding sequentially in time, with reliable temporal relationships embedded in the sequences. Second, the subjects must later be given an explicit test that compares conscious awareness of these relationships with either a forward-directed cue (prediction) or a backward-directed cue (retrodiction).

EXPERIMENT 1

Method

Subjects. A total of 207 University of California at San Diego undergraduates participated in the study.

Apparatus. The experiment was administered on a personal computer with a 19-in. monitor (1,280 × 1,024 pixels).

Stimuli. Eight abstract shapes (Figure 1) modeled after those created by Fiser and Aslin (2001) were used as stimuli. The shapes were black on a white background and scaled to 144 × 144 pixels. From a 1-m viewing distance, each shape measured approximately 2º of visual angle in height and width. Shapes were displayed one at a time, centered on the screen.

The sequence of shapes was determined by a Markov chain—a stochastic process in which the transition to the next state depends only on the current state. Each state is associated with a set of transition probabilities to any of the other possible states. By manipulating these transition probabilities, sequences of shapes exhibiting desired properties (such as predictive relationships) could be constructed by mapping each of the eight shapes onto one of the states in the Markov model, and then allowing the Markov model to determine the transition probabilities between shapes. See Figure 2 for a graphical representation of the Markov model; the matrix of transition probabilities
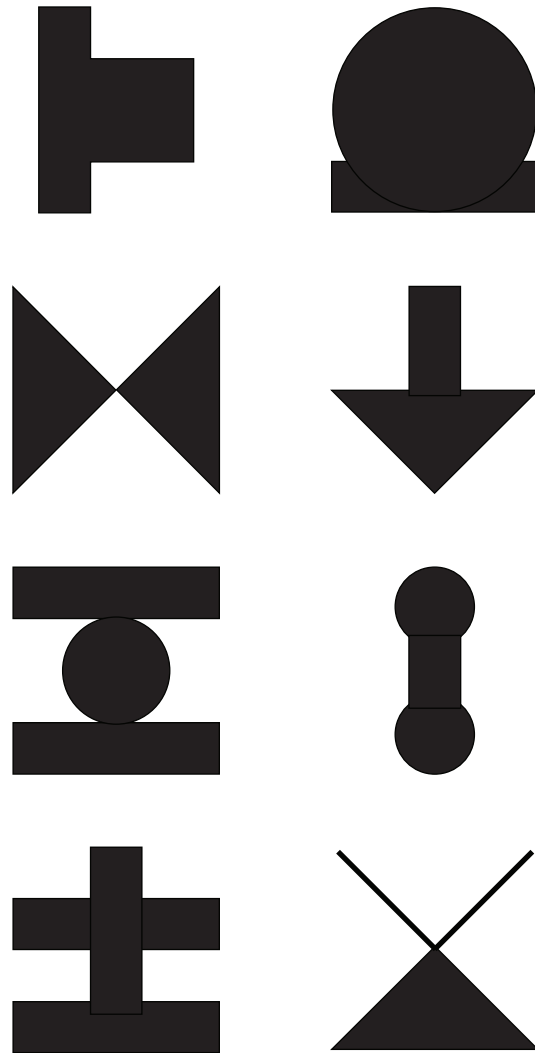

Figure 1. The eight shapes used as stimuli.

used is shown in Table 1. Shapes were randomly mapped to states, independently for each subject. The initial state was randomly chosen.

This set of transition probabilities was devised to create two classes of shapes—those in predictive pairs and those not. Four of the states were organized into two pairs (1–2 and 7–8), in which the first state always predicted the second. In other words, State 1 always transitioned to State 2, and State 7 always transitioned to State 8. The reverse transition was never made; States 2 and 7 were equally likely to transition to any of the nonpaired states or the first state in the opposite predictive pair. The four remaining states (3, 4, 5, and 6) were equally likely to transition to any other state except for 2 and 8 (which were only preceded by 1 and 7, respectively). In terms of shapes, this meant that there were two shapes in the collection that could accurately predict the next shape in the series, whereas all of the other shapes were relatively uninformative as to which shape would follow.

Procedure. The subjects were seated at the computer and asked to read the instructions displayed on the monitor. The subjects were told to attend to a sequence of shapes, but were not advised of the purpose of the study, or about what task would follow. Each shape was displayed for 500 msec, with a 250-msec blank screen intervening between shapes. A total of 600 shapes were presented; the presentation required 7.5 min.
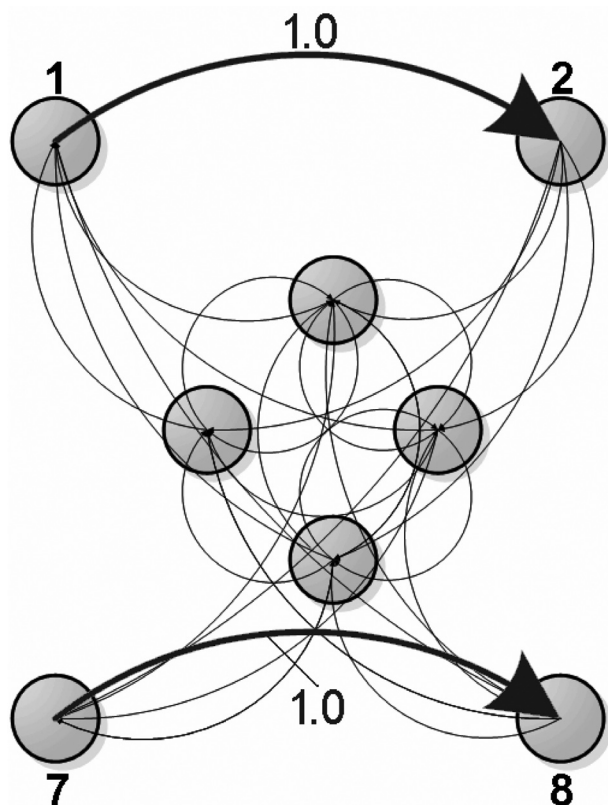
Figure 2. A representation of the Markov model used in this study. Heavy arrows represent the critical predictive relationships and are labeled with the appropriate transition probabilities. Light arrows represent transitions of equal relative probability; the exact values are listed in Table 1. Nonpredictive states (3, 4, 5, 6) are not labeled.

Following this presentation phase, the subjects were informed that they would be tested on the order in which the shapes were presented. Each subject completed a total of two test trials (a prediction question concerning one predictive pair and a retrodiction question concerning the other), with order counterbalanced between subjects. In a prediction test trial, the question "Which shape was most often presented immediately after the shape below?" was displayed at the top of the screen, with all eight shape alternatives presented below. In the middle of the screen (from left to right), the prompt shape, a right-pointing arrow, and a question mark graphic were displayed. Retrodiction test trials were similar except that the subjects were asked which shape preceded the prompt, and the positions of the prompt shape and question mark graphic were reversed. The sub-

jects were directed to click on one answer using the mouse, and to do so as quickly and accurately as possible. The next test trial followed after a 500-msec pause.

Results

Figure 3 shows accuracy for the two tasks. Overall accuracy was 36% (for prediction, 39%, and for retrodiction, 33%). This difference was not significant by a two-tailed Fisher exact test ($p > .25$). The 95% confidence interval for this difference is $6\% \pm 9\%$ (i.e., -3% to 15%).

The mean RT for correct answers was 7,562 msec, $SD = 3,922$ msec, in the prediction condition, and 7,809 msec, $SD = 5,614$ msec, in the retrodiction condition, a nonsignificant difference [$F(1,146) = 0.08$, $p > .75$]. These RTs are higher than those obtained in other probed-recall memory experiments. In the previously cited Kahana and Caplan (2002) study, for example, mean correct RTs inhabited the range of 2,000–3,500 msec, whereas mean RTs in this study were more than double the higher end of that range. We speculate that responses were slower in this study for several reasons. First, subjects completed only two test trials (one in each test direction), rather than a sequence of many trials during which they would become practiced at the task. Second, the subjects did not know the direction of the test until it was presented (and the reaction timing began). Third, the response was made by clicking on one of eight shapes, and thus, making a response involved a visual search and the manipulation of the mouse, rather than a (presumably faster) keypress or verbal response. Another point to note is that no upper limit was placed on RTs, and no data were excluded from analysis. Although the subjects were instructed to respond as quickly and accurately as possible, the range of individual correct response RTs (1.3–28.2 sec) reveals that some subjects responded correctly only after extensive contemplation. Median RTs in the two conditions (prediction, 6,695 msec; retrodiction, 6,575 msec) were not very different from the means reported above, however.

Discussion

Accuracy in the test trials was 36%, well above chance performance of 12.5%. This is consistent with previous findings that people are able to detect differences in transition probabilities in streams of stimuli even when the learning is purely incidental (Aslin, Saffran, & Newport, 1998; Fiser & Aslin, 2002; Hunt & Aslin, 2001). The data

Table 1
Transition Probabilities for Experiments 1 and 2

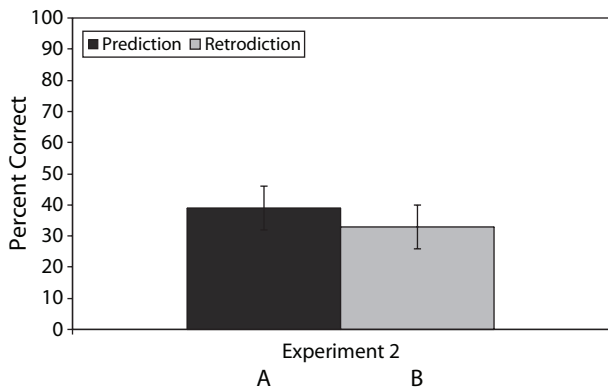| State Transitioning From	State Transitioning To							
	1	2	3	4	5	6	7	8
1	0	1	0	0	0	0	0	0
2	0	0	.2	.2	.2	.2	.2	0
3	.2	0	0	.2	.2	.2	.2	0
4	.2	0	.2	0	.2	.2	.2	0
5	.2	0	.2	.2	0	.2	.2	0
6	.2	0	.2	.2	.2	0	.2	0
7	0	0	0	0	0	0	0	1
8	.2	0	.2	.2	.2	.2	0	0

Figure 3. Accuracy in each test-direction condition in Experiment 1.

provide no evidence that prediction accuracy exceeds retrodiction accuracy, however.

In the next experiment, we examine whether encoding strategies play a role in prediction or retrodiction performance, and obtain further data to compare prediction and retrodiction in general.

EXPERIMENT 2

The previous study suggests that prediction enjoys no sizable advantage over retrodiction. One might still suggest, however, that there is a superior ability to detect forward relationships, but this asymmetric machinery does not operate in purely incidental learning tasks like those explored above. This possibility will be examined in the next experiment, in which some subjects were told to expect a particular type of test. The use of a larger online subject pool also allowed us to follow up on the small but nonsignificant advantage observed for prediction in Experiment 1.

Method
Subjects. A total of 353 members of our laboratory's online subject pool participated in the study, in return for a chance at winning a cash prize.

Procedure. The experiment program was translated to a Macromedia Flash version, which could be administered to subjects online using only their Web browser. The Markov model was the same as that used in Experiment 1. Subjects were randomly assigned to one of three instruction conditions: (1) expect prediction, in which subjects were told to expect a test after the presentation phase, in which they would be given a shape and asked which shape most often followed it; (2) expect retrodiction, in which subjects were told to expect a test in which they would be given a shape and asked which shape most often preceded it; and (3) no expectation, in which subjects were given no information about what sort of test would follow the presentation phase. The no-expectation condition provided an online replication of Experiment 1. In all of the conditions, both prediction and retrodiction test trials were administered as before.

Results
Figure 4 shows accuracy by condition. In the no-expectation condition, the overall percent correct was 38%—very similar to the observed accuracy of 36% in Experiment 1. Again, no significant difference due to test direction was found (prediction, 35% correct; retrodiction, 40% correct, $p > .50$ by two-tailed Fisher's exact test). When these data are combined with the data of Experiment 1, the results show 37% accuracy for prediction and 35% for retrodiction, again not significant by a Fisher's exact test ($p > .65$). The 95% confidence interval for the difference between prediction and retrodiction accuracy for the combined data (2% ± 7%) provides no evidence for any notable superiority of prediction in incidental learning conditions.

Turning to the conditions in which subjects were led to anticipate a particular test, overall accuracy was higher in the expect prediction (44%) and the expect retrodiction conditions (53%) than in the no-expectation condition (38%). The advantage for expect retrodiction was significant [$\chi^2(1, N = 462) = 10.81, p < .001$], whereas the advantage for expect prediction (44%) approached but did not reach significance [$\chi^2(1, N = 484) = 2.29, p = .13$]. Importantly, however, in neither of the expectation conditions was forward recall any better than backward recall. Specifically, in the expect prediction condition, the 95% confidence interval for a potential prediction advantage was −3% ± 12% (or −15% to 9%), and in the expect
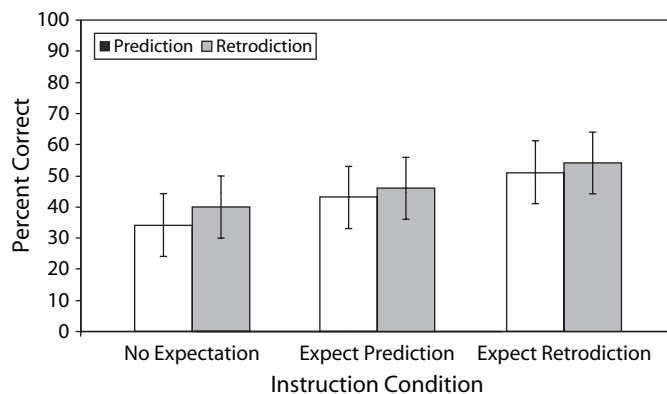

Figure 4. Accuracy in each preexposure instruction condition in Experiment 2, separated by test direction.

retrodiction condition it was −5% ± 13% (or −18% to 8%).

Table 2 presents the RT data for correct responses in each condition. Neither test direction nor expectation condition produced a main effect in RTs, nor were there any significant interactions.

Discussion

Warning subjects in advance that they would be asked to identify a specific temporal relationship between shapes in the presented series aided performance. Interestingly, however, it did not facilitate performance on an expected-direction test any more than it facilitated performance on an unexpected-direction test.

GENERAL DISCUSSION

If human memory is fundamentally specialized for the detection of forward-looking temporal relationships, as theorists of various stripes and from various disciplines have proposed over the past century (Bradley, 1887; Hawkins & Blakeslee, 2004; Llinás, 2001; Rao & Sejnowski, 2003), the experiments described here should have provided an excellent opportunity for this temporal asymmetry to manifest itself. Subjects were reasonably good at learning temporal relationships (even incidentally, as in each of the experiments, as well as intentionally, as in Experiment 2). However, in no case was there any evidence for any temporal asymmetry favoring prediction over retrodiction.

As pointed out by Hoenig and Heisey (2001), a null result is best interpreted in light of the confidence interval for the observed difference between conditions. The combined data for all of the conditions in these studies support an interval for the possible accuracy advantage for prediction ranging between −5% and 6%. This range encompasses the possibility of a 0% advantage (no actual difference between the conditions) and even a small advantage for retrodiction. What it precludes, however, is a large advantage for prediction. These data thus present a challenge to Hawkins and Blakeslee's intuitively very reasonable-sounding suggestions that "our brains use memories to constantly make predictions about everything we see, feel and hear" (p. 86) and "prediction is not just one of the things [the] brain does. It is the primary function of the neocortex, and the foundation of intelligence" (p. 89).

It is interesting to compare the negative findings from the present experiment to Waugh's (1970) observation that repeated practice in forward recall (A→B) led to reduced latencies in subsequent forward recall, without any reduction in latency for backward recall (B→A). Subjects did not practice recalling the observed sequences in the present experiment, but they did experience multiple A to B transitions while viewing the sequences. The absence of a difference between test-direction conditions in accuracy or latency is thus consistent with Waugh's conclusion that it was specifically the effect of recall practice that led to the observed differences in her experiment.

The results obtained here can also be seen as generalizing the findings of associative symmetry observed with intentional verbal paired associate learning tasks (Asch & Ebenholtz, 1962; Kahana, 2002) to a far broader—indeed ubiquitous—kind of situation in which human beings find themselves; namely, experiencing stochastic series of events with pockets of predictability.

Limitations

Nevertheless, a number of limitations of our results should be acknowledged, each of which suggests intriguing directions for further research. As an anonymous reviewer pointed out, in the sequences presented, the two states in each predictive pair were more likely to be temporally adjacent than any other pair of states, regardless of order. Table 3 presents the expected proportion of all transitions in the sequence that will involve the pair of states indicated by the row and column headings. As can be seen, the expected proportion of all transitions that are transitions between States 1 and 2 or between States 7 and 8 (.125) is greater than that for any other pair of states. It may be that subjects are sensitive to the relative frequency with which states are temporally adjacent (without regard to order) rather than the specific transition probability from one state to the next. If that were the case, subjects would only have knowledge that the states in predictive pairs were associated, and not knowledge of the order in which they appeared. They may have relied on this knowledge of association when answering the prediction and retrodiction questions, and the lack of difference between prediction and retrodiction may reflect a lack of knowledge of the transition probabilities ostensibly being tested. It would be enlightening to test subjects for their knowledge of association between states separately from their awareness of the order in which they progressed.

Knowledge of the frequency with which states were temporally adjacent may be independent of knowledge of transition probabilities. Thus, it is unclear whether the results obtained in this study reflect the acquisition of two, equally strong, directional associations for the states in predictive pairs (one forward, one backward) or rather the acquisition of a single association between these states, in which direction is not represented. The results of the present study do not distinguish between these possibilities—they merely demonstrate that memories are not formed in a way that gives an advantage to prediction over retrodiction.

A second limitation is that the experiences that subjects were asked to predict or retrodict lacked any strong hedonic valence. As emphasized by Montague, Hyman,

Table 2
Correct Response Reaction Times (in Milliseconds)
for Experiment 2

| | Expectation | | | | | |
| | None | | Prediction | | Retrodiction | |
Test Direction	M	SD	M	SD	M	SD
Prediction	8,023	5,633	6,742	5,331	7,950	5,049
Retrodiction	8,305	4,752	7,230	6,035	6,923	3,749

Table 3
Expected Proportion of All Transitions for Each State Pair

State	State							
	1	2	3	4	5	6	7	8
1	0	.125	.025	.025	.025	.025	0	.025
2		0	.025	.025	.025	.025	.025	0
3			0	.05	.05	.05	.025	.025
4				0	.05	.05	.025	.025
5					0	.05	.025	.025
6						0	.025	.025
7							0	.125
8								0

and Cohen (2004), it may be that events associated with reward or punishment cause prediction-focused learning mechanisms to become active, whereas affectively neutral events like those used here do not.

The present studies have focused exclusively on *explicit* prediction and retrodiction. It is of course possible that more implicit forms of testing could yet yield evidence for temporal asymmetry. Indeed, it seems plausible that inherently forward-directional representations may be created and strengthened whenever a person repeatedly produces a fixed sequence of motor actions, as suggested by studies of implicit sequence learning (Nissen & Bullemer, 1987) and Waugh's (1970) observation of a reduction in response latency only in the practiced recall direction. Presumably, the common observation that people are better at reciting the alphabet forward than backward also reflects the existence of inherently directional motor plans. A temporal asymmetry confined to sequential motor plans that have been repeatedly performed is, however, a far cry from an overall specialization of the memory system for prediction.

Conclusion

These various limitations notwithstanding, the present results suggest that in response to F. H. Bradley's 1887 question "Why do we remember forwards and not backwards?" we may tentatively answer, "Not so—we are equally good at remembering in either direction." Aside from its inherent interest as a fact about human psychology, this observation may ultimately offer a useful constraint in the development of realistic neurocomputational models of learning and memory.

REFERENCES

ASCH, S. E., & EBENHOLTZ, S. M. (1962). The principle of associative symmetry. *Proceedings of the American Philosophical Society*, **106**, 135-163.

ASLIN, R. N., SAFFRAN, J. R., & NEWPORT, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, **9**, 321-324.

BRADLEY, F. H. (1887). Why do we remember forwards and not backwards? *Mind*, **12**, 579-582.

CHANG, R. C., STOUT, S., & MILLER, R. R. (2004). Comparing excitatory backward and forward conditioning. *Quarterly Journal of Experimental Psychology*, **57B**, 1-23.

FISER, J., & ASLIN, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science*, **12**, 499-504.

FISER, J., & ASLIN, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **28**, 458-467.

HAWKINS, J., & BLAKESLEE, S. (2004). *On intelligence*. New York: Times Books.

HOENIG, J. M. & HEISEY, D. M. (2001). The abuse of power: The pervasive fallacy of power calculations for data analysis. *American Statistician*, **55**, 19-24.

HUNT, R. H., & ASLIN, R. N. (2001). Statistical learning in a serial reaction time task: Access to separable statistical cues by individual learners. *Journal of Experimental Psychology: General*, **130**, 658-680.

JANTZ, E. M., & UNDERWOOD, B. J. (1958). R–S learning as a function of meaningfulness and degree of S–R learning. *Journal of Experimental Psychology*, **56**, 174-179.

KAHANA, M. J. (1996). Associate retrieval processes in free recall. *Memory & Cognition*, **24**, 103-109.

KAHANA, M. J. (2002). Associative symmetry and memory theory. *Memory & Cognition*, **30**, 823-840.

KAHANA, M. J., & CAPLAN, J. B. (2002). Associative asymmetry in probed recall of serial lists. *Memory & Cognition*, **30**, 841-849.

LLINÁS, R. R. (2001). *I of the vortex: From neurons to self*. Cambridge, MA: MIT Press.

MONTAGUE, P. R., HYMAN, S. E., & COHEN, J. D. (2004). Computational roles for dopamine in behavioural control. *Nature*, **431**, 760-767.

MURDOCK, B. B. J. (1962). Direction of recall in short-term memory. *Journal of Verbal Learning & Verbal Behavior*, **1**, 119-124.

NISSEN, M. J., & BULLEMER, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, **19**, 1-32.

RAO, R. P. N., & SEJNOWSKI, T. J. (2003). Self-organizing neural systems based on predictive learning. *Philosophical Transactions of the Royal Society of London: Series A*, **361**, 1149-1175.

SPOONER, A., & KELLOGG, W. N. (1947). The backward conditioning curve. *American Journal of Psychology*, **60**, 321-334.

SUTTON, R. S. (1988). Learning to predict by the methods of temporal differences. *Machine Learning*, **3**, 9-44.

THOMAS, J. G., MILNER, H. R., & HABERLANDT, K. F. (2003). Forward and backward recall: Different response time patterns, same retrieval order. *Psychological Science*, **14**, 169-174.

WAUGH, N. C. (1970). Associative symmetry and recall latencies: A distinction between learning and performance. *Acta Psychologica*, **33**, 326-337.