1997

Long-Term Semantic Priming: A Computational Account and Empirical Evidence

Suzanna Becker
McMaster University

Morris Moscovitch
University of Toronto

Marlene Behrmann
Carnegie Mellon University

Steve Joordens
University of Toronto at Scarborough

Journal of Experimental Psychology:
Learning, Memory, and Cognition
1997, Vol. 23, No. 5, 1059-1082

0278-7393/97/$3.00

Long-Term Semantic Priming: A Computational Account and Empirical Evidence

Suzanna Becker
McMaster University and Rotman Research Institute

Morris Moscovitch
University of Toronto and Rotman Research Institute

Marlene Behrmann
Carnegie Mellon University and Rotman Research Institute

Steve Joordens
University of Toronto at Scarborough

Semantic priming is traditionally viewed as an effect that rapidly decays. A new view of long-term word priming in attractor neural networks is proposed. The model predicts long-term semantic priming under certain conditions. That is, the task must engage semantic-level processing to a sufficient degree. The predictions were confirmed in computer simulations and in 3 experiments. Experiment 1 showed that when target words are each preceded by multiple semantically related primes, there is long-lag priming on a semantic-decision task but not on a lexical-decision task. Experiment 2 replicated the long-term semantic priming effect for semantic decisions with only one prime per target. Experiment 3 demonstrated semantic priming with much longer word lists at lags of 0, 4, and 8 items. These are the first experiments to demonstrate a semantic priming effect spanning many intervening items and lasting much longer than a few seconds.

Many forms of priming have been studied (for reviews, see Monsell, 1985; Richardson-Klavehn & Bjork, 1988; Schacter, 1987). Whereas in repetition priming the priming stimulus is identical to the target, in similarity-based priming tests (e.g., form priming, morphological priming, and semantic priming), the prime and target are different words sharing some surface features, semantic features, or both. Repetition priming and form priming have been found to produce long-lasting effects ranging from hours to weeks or even months (e.g., Bentin & Feldman, 1990; Bentin & Moscovitch, 1988; Jacoby & Dallas, 1981; Rueckl, 1990; Sloman, Hayman, Ohta, Law, & Tulving, 1988). Semantic priming, however, is traditionally thought to produce only short-term effects that dissipate after several seconds or after more than one item intervenes between prime and target stimuli.

Suzanna Becker, Department of Psychology, McMaster University, Hamilton, Ontario, Canada, and Rotman Research Institute, Baycrest Centre for Geriatric Care, North York, Ontario, Canada; Morris Moscovitch, Department of Psychology, University of Toronto, Toronto, Ontario, Canada, and Rotman Research Institute; Marlene Behrmann, Department of Psychology, Carnegie Mellon University, and Rotman Research Institute; and Steve Joordens, Department of Psychology, University of Toronto at Scarborough, Scarborough, Ontario, Canada.

Financial support for this research was provided by Grant 92-40 from the McDonnell-Pew Program in Cognitive Neuroscience, by individual research grants from the Natural Sciences and Engineering Research Council of Canada, and by a Medical Research Council scholarship and Operating Grant MH54246-01. We thank David Plaut for his invaluable comments on this article, Brenda Melo for assistance in data analysis, and Mary McCauley for her assistance in data collection and analysis.

Correspondence concerning this article should be addressed to Suzanna Becker, Department of Psychology, McMaster University, Hamilton, Ontario, Canada L8S 4K1. Electronic mail may be sent via Internet to becker@mcmaster.ca.

Is it possible that completely different priming mechanisms are operating at semantic levels of processing as compared with other levels at which priming could occur? The most parsimonious account would be that the same mechanisms operate at all levels of the system. In this article, we are concerned particularly with long-term priming and argue in favor of a single mechanism to account for all types of long-term priming. Our view is that short-term semantic priming involves a process completely different from that underlying long-term priming, but either type of process should behave according to the same computational principles at any level of the system, whether it be perceptual or semantic. Although our account of long-term priming is very general, our focus is specifically on semantic priming because our model makes novel predictions in this domain. We first present a theoretical account of long-term priming based on a distributed connectionist model of word recognition, combined with some very general learning–processing assumptions. The theory specifies conditions under which long-term priming should occur and predicts that semantic priming should produce long-term effects under the appropriate conditions (even though it has not been found in the literature to date). We use a combination of connectionist modeling and experimental techniques to test our computational account of long-term priming. Whereas computational models have been used most often in the literature to account for data after the fact, our theoretical account has been used to generate predictions and guide our experimental investigations. Our modeling approach is also novel in that we take into account task-specific differences that we predict will explain the failure of previous studies to find long-term semantic priming. Both computer simulations and experiments with human participants are shown to confirm our predictions.

We now review the different forms of long-term priming

that have been studied in the literature and summarize the major theoretical accounts that have been proposed. We then present our computational account of long-term priming and explain why the model predicts that it should be possible to produce long-term semantic priming.

Repetition Priming

Repetition priming is observed when participants are more accurate or efficient in responding to previously studied (primed) targets than to new (unprimed) targets. The term *implicit memory* was coined by Graf and Schacter (1985; Schacter, 1987) to refer to the sort of unconscious memory measured on priming tasks. One of the earliest accounts of long-term repetition priming was based on Morton's (1969) logogen model of word recognition. Morton proposed that long-term priming is the result of a word-detector unit's threshold being lowered as a result of previous suprathreshold activation of that word unit. Bavelier and Jordan (1993) proposed a similar account of long-term repetition priming in a multilayered connectionist model. McClelland and Rumelhart (1986) proposed an alternative mechanism for long-term repetition priming in neural networks with recurrent connections, which we adopt in our model. They postulated that exposure to each pattern involves some incremental learning. (See also Bower, 1996, for a similar account.) We refer to this postulate as the "incremental learning hypothesis." It predicts that all of the connections in the network that are involved in processing a pattern, not just the thresholds of units, should undergo some incremental learning as a result of priming. In McClelland and Rumelhart's model, each learning step involves a large initial change in each weight, which rapidly decays down to a permanent or very slowly decaying smaller change. Thus, priming is thought to reflect the normal course of learning. When a network is exposed to a previously primed pattern, it should settle to a stable response more quickly because the connections involved in producing the response have been reinforced.

Similarity-Based Priming

The threshold lowering and incremental learning hypotheses both can account for long-term repetition priming, in which responses to the same input pattern are faster or more accurate on repeated presentations. However, when the prime and target stimuli are not identical but related, the threshold account predicts no long-term effect. The incremental learning account makes no specific prediction about how long-term priming would depend on the level of similarity between primes and targets, although McClelland and Rumelhart (1986) did make the general argument that long-term priming should generalize from primes to similar targets on the basis of the amount of overlap in their representations. For example, in form-based priming (Forster, 1987), the prime and the target are similar in that they share perceptual features but are otherwise unrelated, whereas in semantic priming the prime and the target are semantically similar or are semantic or contextual associates. Any theory capable of making specific predictions about these cases must make stronger assumptions about the nature of the representation, the processing of stimuli, or both.

Most of the experimental work on form-based priming has examined short-term effects, as first reported by Meyer, Schvaneveldt, and Ruddy (1974), who found faster lexical decision for words preceded by orthographically and phonologically similar words. In contrast, long-term similarity-based priming has been much less studied. Bentin and Feldman (1990) reported that Hebrew words based on a common morphological root, that is, the same consonant pattern, produced a priming effect on lexical decision that was still significant with 15 intervening items between the prime and the target. This was true regardless of whether or not the words were semantically associated. When words were semantically and not morphologically related, the priming effect fully dissipated after a lag of 15 items. Rueckl (1990) has reported consistent long-term priming effects in tachistoscopic word recognition, when each word was primed by a large set of orthographically similar words. Studies by Rueckl and Olds (1993) and by Feustel, Shiffrin, and Salasoo (1983) provided converging evidence for these results. Rueckl interpreted the results of his 1990 studies by using McClelland and Rumelhart's (1986) incremental learning hypothesis, combined with an assumption about the distributed nature of word representations. It was predicted that the amount of form priming should increase with the similarity between prime and target. Assuming an orthographic representation of words in which activity was distributed across many units, Rueckl reasoned that orthographically similar words would activate many features in common and therefore would benefit mutually from the strengthening of connections between those features. Note that Rueckl's findings cannot be explained adequately by theories that postulate long-term threshold changes in word-detector units; these "localist" theories would not predict such effects to generalize to visually similar words.

Whereas repetition and form priming effects may be very long lasting, semantic priming (Meyer & Schvaneveldt, 1971) on implicit memory tests such as lexical decision and naming has been found to disappear when the prime and target are separated by several intervening items or time lags of more than several seconds (e.g. Bentin & Feldman, 1990; Dannenbring & Briand, 1982; Henderson, Wallis, & Knight, 1984; McNamara, 1992b; Monsell, 1985; Ratcliff, Hockley, & McKoon, 1985). Studies of semantic priming across a lag of a single intervening item (reviewed in Joordens & Besner, 1992; Masson, 1995) have yielded mixed results; Joordens and Besner (1992) have shown a very small but reliable semantic priming effect in the lexical-decision task across a lag of a single item, even when the possibility of conscious comparison of primes and targets is minimized. Nevertheless, the above literature would suggest that semantic priming has little or no effect beyond the immediate semantic context.

We now describe our view of long-term priming in connectionist networks, which predicts that a long-term semantic priming effect should also be possible. To anticipate, we propose that long-term semantic priming involves

incremental learning in semantic networks rather than a process such as residual activation that mediates short-term effects. Thus, our proposed model of long-term semantic priming is not meant, for the moment at least, to rival models that seek to account for short-term priming. Though fundamentally different from each other, the processes underlying short- and long-term semantic priming are seen as complementary. One is likely to be crucial for supporting the type of rapid, but transient, anticipations necessary for efficient discourse and reading, whereas the other reflects and contributes to the construction of long-lasting semantic networks. As a result, the type of test that best reveals long-term semantic priming will be different from the one that is sensitive to short-term priming. The overwhelming majority of short-term priming tests have focused on the lexical level, whereas for long-term priming to be revealed, the tests need to access the semantic level to a greater degree. These issues are addressed at greater length in presenting our model and the experiments that support it.

Word Recognition in Attractor Networks

To account for differences in similarity priming when the primes are visually versus semantically related within a computational framework, we begin by describing a connectionist model of word recognition that involves multiple levels of processing. This model was originally proposed by Hinton and Shallice (1991) to account for symptoms of reading impairment such as those seen in people with deep dyslexia. The network was trained with an iterative version of the back-propagation learning procedure (Rumelhart, Hinton, & Williams, 1986) to map orthographic representations of words onto distributed semantic representations via a layer of hidden units, with the help of semantic "cleanup units"—an extra, hidden layer only connected to the semantic units. There were feed-forward connections from orthographic to hidden units and reciprocal connections between the hidden and semantic units, between the cleanup and semantic units, as well as within the semantic layer.

Using this architecture, Hinton and Shallice (1991) proposed a novel view of the mechanism underlying word recognition: When a word is presented to the network as an orthographic input pattern, the rest of the network, including the semantic and hidden layers, gradually settles into a stable representation called an *attractor state.* Recognition occurs when it reaches the correct attractor. The current state vector of the input units can be thought of as representing a single point in a high-dimensional orthographic space, with dimensionality equal to the number of orthographic input units. Likewise, the current state vector of the semantic units can be thought of as a single point in semantic space, with dimensionality equal to the number of semantic units. Because there is no systematic correspondence between orthographic and semantic features, the network must learn to map points that may be nearby in the orthographic state space onto much more distant points in the semantic state space. In contrast, semantically related words are typically represented by distant points in orthographic space, but by nearby points in semantic space because they share many semantic features. When the network is damaged by removing connections or by adding noise to the weights, the reading pattern associated with deep dyslexia is observed: Not only are visually similar words (e.g., *cat* and *mat*) seen to be occasionally confused, but substitution errors involving semantically related words are also observed (e.g., *cat* read as *dog*), as are mixed visual–semantic errors (e.g., *cat* read as *rat*).

Long-Term Semantic Priming in Attractor Networks

Adopting the Hinton–Shallice (1991) view of word recognition leads us to conceptualize the process of long-term priming rather differently. When a word is presented to the network, and its semantic representation is activated, assuming McClelland and Rumelhart's (1986) view of long-term priming is correct, all the connections participating in the entire activated pathway of units should be altered. These weight changes should increase the probability that the network will produce the same response when given the same input pattern in the future, even if the pattern is noisy or incomplete, and should make the same units even more strongly active the next time (if they are not already at their maximum activation levels). Thus, the attractor for this particular input pattern will be deepened.[1] Because the shape of the attractor basin has now been altered, the network's responses to other patterns should also be affected. Small perturbations in the state of the network in the vicinity of this attractor should be pulled more strongly back to the attractor state. Semantically related words might therefore be expected to benefit from their nearby neighbor having been primed because their attractors would overlap in many dimensions. Priming with multiple semantically similar primes should produce an even larger effect than with single-word semantic priming. Balota and Paul (1996) found empirical support for this prediction in the context of short-term priming: Two primes produced more priming than did one prime. The computational reason for this prediction in the case of long-term priming is that the semantic attractors for the primes would be deepened preferentially along those dimensions common to many elements of the prime set and common to the target as well. However, those dimensions of the attractor basins for the primes that were not shared by the target would be relatively unaffected.

To make the above predictions more concrete, if we could measure at the semantic level the settling time of the network, we would expect to see long-term semantic priming. However, because semantically related words normally are far apart in the orthographic space, a semantic prime would not be expected to influence settling time to the same degree at the orthographic level. Similar predictions

[1] Plaut and Shallice (1993b) have proposed a similar scheme to account for perseverative naming in optic aphasic patients, which they described as being analogous to a temporary priming effect. They implemented this priming effect by using a separate set of fast-decaying weights to store short-term correlations between units' activities across recently presented patterns.

could be made regarding human participants: When primed with a word having similar semantics to the target, participants should be faster on a semantic-retrieval task such as semantic classification. On tasks such as lexical decision or word identification that depend less directly on semantics, however, there should not be strong evidence of semantic priming.

Why have previous experiments failed to show long-term semantic priming? Our predictions suggest that the following two variables are critical: (a) the task and, in particular, the level of processing in which the participant engages when responding to both prime and target words, and (b) the degree of semantic overlap, or number of shared features, between the primes and the targets. To our knowledge, with the exception of Woltz's (1990, 1996) work (see General Discussion), all of the previous attempts to study the time course of semantic priming have used word-recognition tasks such as lexical decision rather than explicitly semantic tasks such as animacy or size decisions. Our prediction is that the semantic priming effect occurs by deepening attractors in semantic space for both primes and related targets, and that this effect should primarily manifest itself on semantic-retrieval tasks. Further, a close match between the type of processing involved during presentation of the prime and target should also be critical. This last point has been demonstrated previously in a series of priming experiments by Vriezen, Moscovitch, and Bellos (1995). Except for one experiment, these studies assessed *repetition priming* in a semantic-classification task. Vriezen et al. found that when lexical decision or naming on the priming word list was followed by a semantic decision (e.g., of size or animacy) on the target word list, there was no priming effect. Only when the task performed on the priming list tapped into the same level of processing (or higher), as compared with the task performed on the target list, did priming occur, although the tasks did not have to be identical.

Our second prediction is that the degree of semantic overlap should also be a critical variable. Only when the prime and target have sufficiently overlapping basins of attraction in semantic space would we predict long-term semantic priming. Thus, words that are highly associated on the basis of free-association norms but are not particularly similar in meaning, such as *bread* and *butter,* would not be expected to produce long-term priming. This prediction makes intuitive sense when priming is viewed as a reflection of incremental learning. We would like representations in memory to reflect our recent experiences and for learning to generalize to items that have overlapping meanings; however, more distant associations such as *bread–butter* might be modeled better by some alternate learning mechanism.

The following simulations and experiments were carried out to test our predictions about long-term semantic priming. Simulation 1 involved semantic priming in a neural network using blocked priming trials, with priming blocks consisting of multiple semantic primes per target. Settling time was the dependent measure. Simulation 2 involved a similar testing procedure, but with only a single related semantic prime per target word. Our previous simulations (S. Becker, Behrmann, & Moscovitch, 1993) investigated repetition priming, form priming and semantic priming, using single-word priming trials, and found an interaction between the prime type (orthographic vs. semantic) and the level of processing during the test (orthographic vs. semantic layer settling time). We designed the current simulations to replicate our previous results on semantic priming. However, we modified our procedure in two substantial ways to simulate more realistically a long-term priming procedure that could be applied to human participants. First, rather than presenting a single related prime immediately preceding each target, we switched to a blocked priming procedure; in each block of trials, several related and unrelated primes were presented, followed by a target word. This is a more realistic simulation of long-term priming because there are multiple intervening items between a related prime and target. Second, we switched the semantic task from semantic layer settling time to a more specific animacy-decision task that could be performed by human participants. In the model, we simulated the animacy decision by measuring the settling time of the unit in the semantic layer representing the animacy feature. In Experiment 1, we measured long-term semantic priming in participants by using blocked priming trials on a semantic-decision (animacy) versus a lexical-decision task, with multiple semantic primes for each target, as in Simulation 1. In Experiment 2, we further investigated the semantic priming effect by using the same experimental setup and target words but only a single word prime per target as in Simulation 2. Experiment 3 was conducted to determine whether long-term priming could be sustained over much longer lists and to explore its decay function by varying the lag between the prime and the target.

Simulation 1

We used a network that we had first trained to perform visual word recognition by producing the correct semantic and phonological representations in response to an input orthographic representation. The phonological layer was not needed for the present simulations but was included in the network for use in other simulations. The network was tested on a series of priming trials, on orthographic- and semantic-decision tasks.

Method

The network architecture. The network used in our simulations is shown in Figure 1. It is similar to Hinton and Shallice's (1991) network (described above), except that we used a deterministic Boltzmann machine (DBM; Peterson & Anderson, 1987) rather than a back-propagation network. We chose the DBM network because it is trained by using a simple contrastive Hebbian learning rule (see Appendix A) that is thought to be more biologically plausible than back-propagation networks. Plaut and Shallice (1993a) showed that a DBM network very similar to ours could produce qualitatively similar behavior to Hinton and Shallice's network. Our network consisted of orthographic, semantic, and phonological (O, S, and P) layers, a hidden layer between the O and S units, and a hidden layer between the S and P units, as shown in Figure 1. The P layer was not relevant to the simulations reported here. As in all DBM networks, all connections were bidirectional.

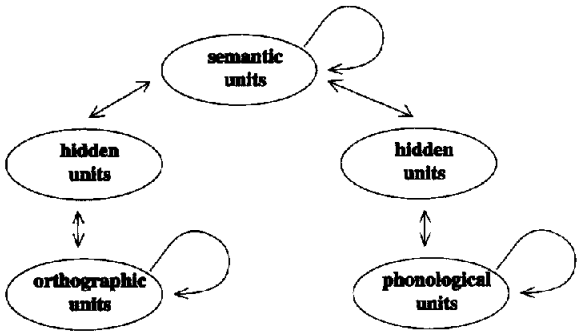

Figure 1. The architecture of the network. Arrows indicate full connectivity among units within or between groups. All connections are bidirectional. There were 40 units in each hidden layer, 28 orthographic units, 33 phonological units, and 68 semantic units.

The feedback connections greatly increase the settling time of the network, particularly because the network must be simulated on a serial machine. To minimize the number of connections in our network, we only included within-layer or autoassociative connections in the O, S, and P layers that were involved in generating responses. It was thought that such connections would enhance within-layer priming effects.

Although our network differs from the Hinton–Shallice (1991) architecture in that it lacks the extra hidden layer of semantic cleanup units, the bidirectional links add enough extra connections to make the cleanup layer unnecessary. Further, the S → P route may provide additional cleanup. Our network architecture was virtually identical to that used by Plaut and Shallice (1993a, Figure 12) except that each unit in our O and P layers also had a link to itself, and each O layer unit had an external input connection. These two modifications permitted us to measure settling time in the O layer.

The unit activations. To be able to study the network's orthographic settling time, we used a slightly different mode of pattern presentation from the standard DBM during the latter stages of training and in the priming simulations. In the standard DBM, an input pattern is presented to the network by fixing the states of the input units to the pattern values and by not allowing these states to update. This method of pattern presentation at the input layer is sometimes referred to as "hard-clamping," and means that settling time cannot be measured at the input layer. We therefore used a "soft-clamping procedure" for the orthographic layer input units, which allowed the orthographic layer to settle along with the rest of the network. In our soft-clamping procedure, orthographic units' states were strongly influenced by their external input but were also subject to top-down influences. Further details on the unit activations and training procedure are provided in Appendix A.

The training procedure. DBM learning (Peterson & Anderson, 1987) is appealing because it is based on a simple Hebb-like learning rule. Hebb (1949) postulated that the strength of the connection between two neurons should increase whenever those neurons are simultaneously active. DBM learning is more general in that connection strengths are adjusted in proportion to the product of the pre- and postsynaptic activities, and can therefore either increase or decrease.[2] Learning in a DBM proceeds in two phases, applied alternatingly for each training pattern: (a) a positive phase, in which the input and output units are clamped to their correct states and positive Hebbian learning occurs, and (b) a negative phase, in which only the input units are clamped and negative Hebbian learning or "unlearning" occurs. See Appendix A for further details.

The network was trained for 3,500 sweeps through the training set of 40 patterns until the network learned to produce the correct semantic and phonological representations in response to each orthographic input pattern. No priming effects were simulated during the training phase. Units' states were considered correct when they were within 0.3 of their target states.

Materials. The training set consisted of the same 40 words' orthographic-input and semantic-output vectors used by Hinton and Shallice (1991), shown in Appendixes B and C, augmented by the phonological output vectors used by Plaut and Shallice (1993a). The words were composed of only the letters {*b, c, d, g, h, l, m, n, p, r, t*} in the first position, {*a, e, i, o, u*} in the second position, {*b, c, d, g, k, m, n, p, r, t, w*} in the third position, and {*e, k,* –} in the fourth position, where – denotes a blank. Thus, there was one orthographic input unit for each letter–position combination. Similarly, Plaut and Shallice's (1993a) phonological output representation consisted of 33 position-specific phonemic features: {*/b/, /k/, /d/, /dy/, /j/, /g/, /h/, /l/, /m/, /n/, /p/, /r/, /t/*} in the first position, {*/a/, /e/, /i/, /o/, /u/, /ie/, /ew/, /ar/, /aw/, /ow/, /oa/*} in the second position, and {*/b/, /d/, /g/, /k/, /n/, /m/, /t/,* –} in the third position. Words were grouped into five categories: indoor objects, animals, body parts, foods, and outdoor objects. The Hinton–Shallice semantic features were chosen so that words in the same category had a much higher degree of semantic overlap (80.5% on average) than did words in different categories (66.2% on average).

Procedure. The priming condition for each target word was either unrelated or related, and the dependent measures were orthographic-level and semantic-level settling times. The unrelated primes condition served as a baseline. The network was tested on blocks of 11 words. Each block consisted of a priming list of 10 words, followed by a single target word. Each prime list was constructed so that either half of the words in the list were related to the target (the related primes condition) or all 10 words were unrelated to the target (the unrelated primes condition). The words in each block were randomly chosen without replacement as follows. In the related primes condition, the 5 related words were drawn from the same category as the target word but excluding the target word, and the 5 unrelated words were drawn from one other randomly chosen category excluding the target category. In the unrelated primes condition, two categories were selected at random, without replacement, excluding the target category; 5 words were then randomly selected from each of the two categories. The priming words were presented in a random order, with the constraint that the last 3 prime words must be unrelated to the target.

Priming was simulated as an incremental learning mechanism (McClelland & Rumelhart, 1986). For each prime word in a block, the weights in the network were updated according to the DBM contrastive Hebbian learning rule (see Appendix A), with a fixed learning rate of 0.0002. When all of the prime words in a block had been presented, the network was then presented with a target word and again permitted to settle. To keep the number of free parameters in the model to a minimum, no decay in the learning was simulated. However, to obtain a more realistic model, we would obviously need to address this issue.

Each target word was tested for 20 successive blocks in each of the two priming conditions (unrelated and related primes), with the particular primes selected randomly with replacement in each

[2] The terms *presynaptic* and *postsynaptic,* borrowed from neurobiology, are standard in the connectionist modeling literature. They refer respectively to the units at the sending and receiving ends of a connection.

Table 1
Mean Settling Times (STs) and Error Rates (in Percentages) for Simulations 1 and 2

	Network performance measure					
	Orthographic ST		Semantic ST		Errors	
Simulation	*M*	*SE*	*M*	*SE*	%	*SE*
Simulation 1						
Unrelated	23.16	2.0	33.04	2.1	10	3.6
Related	22.03	1.0	32.03	1.9	5	1.7
Simulation 2						
Unrelated	22.73	1.9	33.87	2.5	9	3.7
Related	22.62	1.8	33.27	2.4	9	3.9

block. To prevent identity priming effects when the same target words were retested in different blocks, we reset the weights in the network after each block. Each time a target word was presented, a small amount of noise with mean 0.0 and standard deviation 0.05 was added to the input (orthographic) units' states, as during training. This repeated testing of each word was done to generate a distribution of independent responses to each target word that could be statistically analyzed.

We used two response measures on each trial, orthographic layer settling time and semantic (animacy) decision time. The orthographic settling time was assumed to be analogous to lexical-decision time. The semantic decision was whether a given word represented an animate or an inanimate object. In our network, as in the Hinton–Shallice (1991) model, the semantic representation included a single unit designated to represent the "animacy" feature. We therefore simply used the settling time for this animacy unit as a measure of the semantic-decision time. A group of one or more units was considered to have settled to a correct response when each of those units was within 0.3 of its correct value, the same correctness criterion used during the learning phase. Because we measured responses from all of the orthographic layer units as well as from the animacy unit on each trial, all of these units were required to be in correct states for the responses on this trial to be considered correct. If the network did not arrive at a correct response by the time all units' states had stopped changing (to within a tolerance of 0.001), it was considered an error trial. Priming scores for correct trials were calculated for each word as the average difference between settling time (ST) on related trials and unrelated trials. The baseline or "unrelated" ST was taken to be the mean settling time for each target word when primed by a set of unrelated words. The "related" ST was measured for each target primed with a mixture of related and unrelated words.

Results and Discussion

The mean STs averaged across words on correct trials and percentages of error responses in each condition for Simulation 1 are shown in the top half of Table 1.[3] On average, the network settled about 1 cycle faster for related primes compared with unrelated primes, at both the orthographic and semantic levels. These priming effects are plotted in the left half of Figure 2. Paired *t* tests on the mean STs for each word on related correct trials versus unrelated correct trials revealed that on the semantic task, priming was significantly greater than zero, $t(39) = 2.37$, whereas on the orthographic task it was not. For these and all subsequent analyses, effects are described as significant if the probability of observing the effect is less than .05; all *t* tests and sign tests were one-tailed. There was a very strong trend toward positive priming on the orthographic ST scores. However, a closer inspection of the mean ST scores for individual words revealed that this trend was not at all consistent across words. At the orthographic level, priming was actually negative for 20 of the 40 words, zero for 4 words, and positive for 16 words. Further, of the 16 positive cases, only 2 were very large and the rest were close to zero. Thus the apparent trend appears to be a result of a couple of outliers. The high variability in orthographic ST scores may be a result of the fact that we used a small set of short words with relatively high orthographic overlap. In contrast, at the semantic level, priming was much more consistent across words. It was negative for only 6 words, zero for 10 words, and positive for 24 words. A paired *t* test on the mean priming scores (unrelated ST minus related ST) for each word at the orthographic level versus the semantic level revealed that the difference was not significant. Again, this appears to be a result of the large variability in the orthographic ST scores. A paired *t* test on the mean error rates for each word on related and unrelated trials revealed no significant differences. Thus, the results are consistent with our prediction that there should be a strong long-term semantic priming effect on a semantic-decision task but not on an orthographic task.

Simulation 2

In Simulation 1 we used multiple related primes per target because it was thought that multiple primes would be most likely to produce long-term priming in human participants. Simulation 2 was run to test whether long-term priming would also be observed with a single related prime per target. It was predicted that there would be less semantic priming in this case.

Method

The same network, tasks, and procedure as in Simulation 1 were used. This time, however, each target word was only primed with a single related prime per block. Twenty blocks were run for each of the 40 target words. Each block once again consisted of a priming list of 10 words, followed by a single target word. In the related primes condition, a block consisted of a word randomly chosen from the same category but not identical to the target, and 9 randomly chosen primes, each from a different randomly selected unrelated category, followed by a target word. In the unrelated primes condition, a block consisted of 10 randomly chosen unrelated primes, each selected from a different one of 10 randomly chosen unrelated categories, followed by a target word. In both conditions, the randomly chosen categories were sampled with replacement. As in Simulation 1, the priming words were presented in a random order, with the constraint that the last 3 prime words must be unrelated to the target.

[3] We used the same error measure for both tasks. Alternatively, one could measure error rates separately for the two layers.

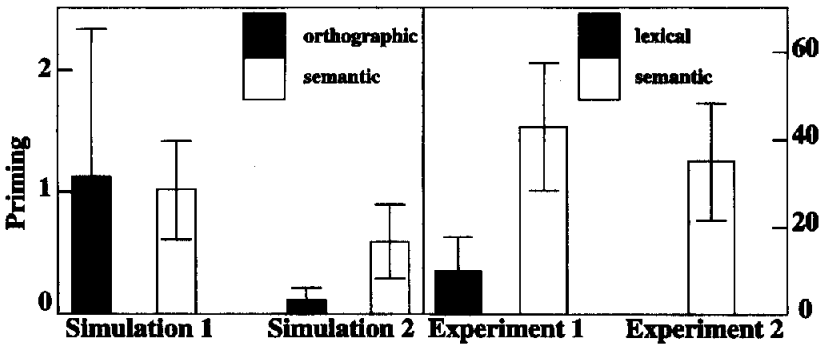

Figure 2. Plot of the unrelated minus the related settling times (cycles) for the orthographic and semantic layers in Simulations 1 and 2, and the unrelated minus the related reaction times (in milliseconds) for lexical and animacy decisions in Experiments 1 and 2. Each bar represents the standard error of the priming effect.

Results and Discussion

The mean STs averaged across words and percentages of error responses in each condition for Simulation 2 are shown in the bottom half of Table 1. On average, the network settled 0.12 cycles faster for related targets at the orthographic level and 0.60 cycles faster for related targets at the semantic level. Paired *t* tests on the mean STs for each word on related and unrelated trials revealed significant priming at the semantic level, $t(39) = 1.93$, but not at the orthographic level. These priming effects are plotted in the left half of Figure 2. An examination of the mean priming scores for individual words revealed that at the orthographic level, semantic priming was positive for 2 words, negative for 1 word, and zero for the rest. At the semantic level, 12 scores were positive, 5 were negative, and the rest were zero. A paired *t* test on the mean priming scores (unrelated STs minus related STs) for each word at the semantic level versus the orthographic level revealed that there was significantly more priming on the semantic task than on the orthographic task, $t(39) = 1.92$. Paired *t* tests on the mean error rates for each word on related and unrelated trials revealed no significant differences at any level. To compare the results of Simulations 1 and 2, we performed a paired *t* test on the mean semantic priming scores for each word. It was predicted that there would be greater semantic priming in Simulation 1 than in Simulation 2. Although there was a trend in this direction, a *t* test revealed that the difference was not significant.

To test more directly our hypothesis that priming involves deepening the attractor basins for primes and related words, we also computed the depth of the attractor basins for words before and after priming. The depth of any point in the state space of a DBM network can be determined by computing the *free energy* of the network (see Hinton, 1989). Details of this computation are provided in Appendix A. Conceptually, the energy can be thought of as a measure of the amount of tension in the network. If a pair of units are strongly linked but they are in opposite states, there is tension between them, raising the energy of the system. The same is true if they are negatively linked but in the same state. Thus, a network in a low-energy configuration has each pair of connected units' states in good agreement with the constraints imposed by the connection weights. The energy is a continuous function over the entire space of possible states of the network. If we could plot this function before and after priming, we should be able to see how the attractors are reshaped. However, the state space is high-dimensional, so it is impossible to display the attractor basins for words graphically. However, we can plot one-dimensional subregions of this state space by computing the energy of any two points in the space, as well as a series of intermediary points along a linear trajectory connecting those two endpoints, to visualize how priming shapes the attractor basins. We start by computing the energy of the two end points, for example, the attractor states for *cat* and *dog*. We then compute the vectorial difference between those two states; this gives us the direction one must travel in state space to go directly from *cat* to *dog*. We can now interpolate points along this path by taking the state vector for *cat* and adding to it some proportion of the difference vector. Setting the network states to these interpolated points, we can compute the energies at each of these points to get a function of energy along this one-dimensional path in state space.

We computed the energy of the network after it had been presented with a prime word, after it had been presented with a target word, and also the energy of 100 states computed at evenly spaced intervals between these two end states, before and after priming. We found that the energy changes were too small to be visible on the graph (usually in the fourth or fifth significant decimal place), although they were highly reliable. We also examined the changes in energy at the semantic and orthographic levels separately. Semantic-level energy was computed by applying the free energy equation given in Appendix A to the units in the semantic layer and by considering only within-layer connections. Orthographic layer energy was computed similarly for the orthographic units. Paired *t* tests were performed on both the means and the medians of these energy scores for each word in the related versus unrelated conditions. In Simulation 1, 34/40 words showed decreases in mean semantic layer energy after priming with related words compared with

the unrelated baseline, and 38/40 words showed decreases in mean semantic layer energy. Although the effect on the mean semantic energies was not significantly different from zero (average unrelated minus related mean semantic energy = .02, standard error = .28, $t[39] < 1$), it was significant for the medians (average of median semantic energy changes = −.44, standard error = .25, = 1.80). The lack of significance in the means was clearly due to 5 outliers that represented huge increases in mean energy (and one small increase), relative to the other 34 words that all showed small but consistent decreases. The pattern of energy changes for the orthographic layer was very similar but in the opposite direction: 29/40 words showed increases in mean orthographic layer energy in the related condition, and 34/40 of the words showed increases for the medians. *T* tests of the orthographic layer energy changes were not significant for either the means or the medians.

Because of these small differences, we therefore used a much larger learning rate of 0.02 rather than 0.0002 for generating the energy plots, which produced more noticeable differences. Typical examples of such graphs are shown in Figure 3 for two pairs of words. In Figure 3A, we see energy plots for a pair of semantically unrelated words (*bed* and *cat*) before and after priming by the word *bed.* The energy at the left edge of the curve, at the attractor state of the word *bed,* is lower after priming with *bed* (repetition priming). The energy of the curve at the rightmost point, at the attractor state for the word *cat,* does not change much if at all after priming with *bed.* Thus, there is no evidence of semantic priming for unrelated words. In Figure 3B, we see energy plots for a pair of semantically related words (*dog* and *cat*) before and after priming by the word *dog.* This time, the primed curve is lower at all points along the trajectory through the state space from *dog* to *cat.* Thus, the basins of attraction have been deepened for both the prime and the target word. In addition, the energy barrier between the two attractor states, that is, the energy maximum near the midpoint of the curve, has been lowered after priming.

We have demonstrated a general pattern of results in our simulations that was predicted by our computational model. Priming deepens the attractor basins of both primes and semantically similar words, leading to faster semantic but not orthographic decisions. Further, we saw a trend toward multiple primes being more effective than single primes at speeding performance on semantic decisions. We view these results as a preliminary test of the feasibility of our model, providing grounds to investigate the long-term semantic priming phenomenon further. The next question is whether similar results would hold in human participants. If the model has captured the essential aspects of semantic memory representation for the current tasks, as well as the basic mechanism of long-term priming, we should see at least qualitative agreement between our pattern of simulation results and the pattern of priming in human participants.

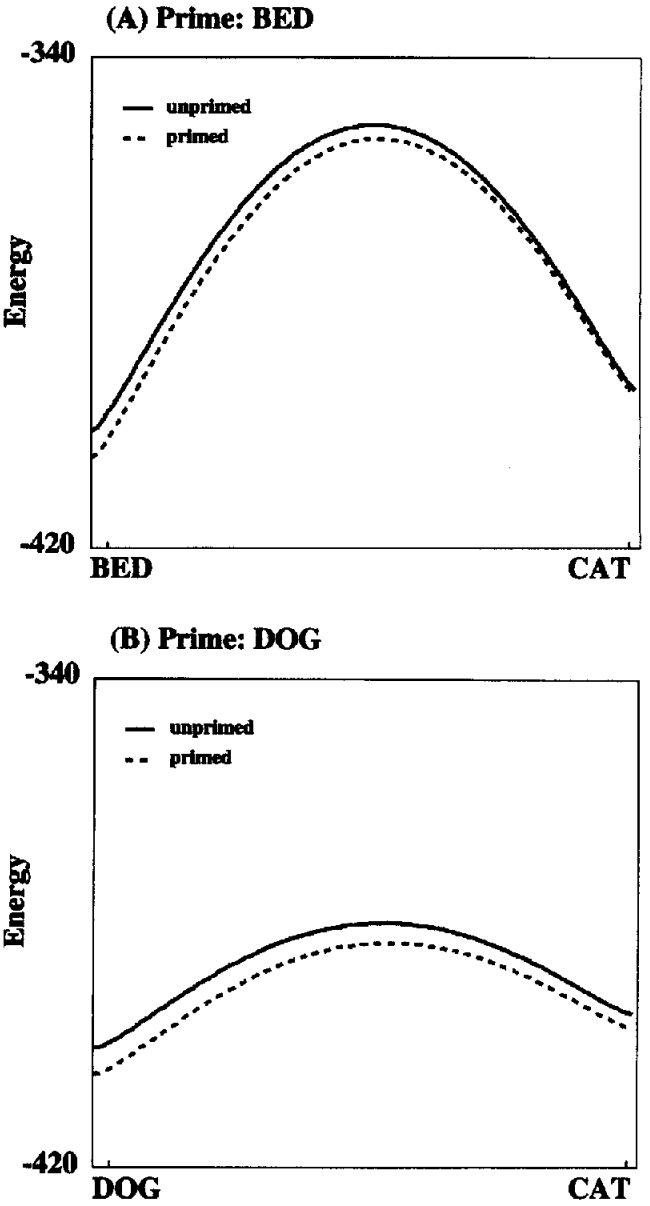

Figure 3. The free energy is plotted against the state of the network before and after priming with (A) an unrelated word and (B) a related word. Each point along the *x*-axis represents a single point in the high-dimensional network state space. The two end points on the *x*-axis of each graph represent the state the network is in when it recognizes the words shown. The intermediary points represent the states of the network along a direct path between the two end states. The distance along the *y*-axis represents the energy. The difference between the upper and lower curves represents the priming effect for repetition priming (left end) and semantic priming (right end).

Experiment 1

On the basis of the results of Simulation 1, we designed Experiment 1 to test the hypothesis that long-term semantic priming in human participants would produce more facilitation on a semantic task than on a lexical-decision task. To keep the conditions as close as possible to the simulation experiments, we attempted to use similar tasks. There is some controversy in the literature as to what performance index in a connectionist network constitutes a good model of

lexical decision. Some have argued that lexical decision could be based, at least in part, on phonology (Rueckl, 1995; Seidenberg & McClelland, 1989). Here, we used orthographic ST to model lexical-decision time in human participants. Seidenberg and McClelland (1989) used orthographic error scores in their networks to model human lexical-decision errors. However, Besner, Twilley, McCann, and Seergobin (1990) found that this model generated unrealistically high error rates for nonword processing, calling into question whether it was even capable of handling nonword performance.[4] Here, we are concerned with reaction times to words rather than to nonword performance in our model. In this respect, as our experiments show, orthographic ST in our model seems to capture human performance in semantic priming quite well, at least qualitatively. Note that we are not claiming that lexical decision is a purely orthographic task; it is likely to be influenced to some degree by semantics. Likewise, in our network simulations, the orthographic layer was influenced by top-down feedback from semantics, although we referred to orthographic ST as an "orthographic task."

For the semantic-decision task, as in the network simulations, the task was an animacy decision; participants judged words to be living or nonliving. For each participant, half of the target words were primed with a set of semantically similar words. Participants performed the same decision for both prime and target words.

Method

Participants. Sixty undergraduate students from the University of Toronto participated in our study for credit in an introductory psychology course.

Materials. The prime words presented to each participant were drawn from a list of 150 words, each chosen to be closely semantically related to 1 of the 30 target words, so there were 5 primes per target. We used multiple primes per target to maximize the chance of obtaining a long-term semantic priming effect, as Rueckl (1990) did in studying long-term form priming. The primes were chosen from the same Battig–Montague category norms (Battig & Montague, 1969) as well as from a thesaurus of synonyms.

The target list consisted of 45 four- and five-letter words. Thirty of these were target words (15 primed and 15 unprimed) and 15 were filler words. To minimize the possibility of any short-term semantic priming from occurring within the target list, we chose each target word from a different category. The fillers were 15 pronounceable nonwords if the task was lexical decision and 15 words that were semantically unrelated to the targets if the task was semantic decision.

The prime words and corresponding targets were divided into two equal-sized sublists, shown in Appendix D. Each participant was presented with only half of the primes in the study list (75 of the 150 primes), either Sublist A or B, and tested on the entire target list (30 words), including both primed and unprimed targets. The sublist of primed targets was counterbalanced across participants, so that each target word was primed for half of the participants and unprimed for the other half.

Procedure. Participants were assigned to either the lexical-decision or the semantic-decision group. Within each of those groups, half were primed on Sublist A and half on Sublist B (Appendix D). Stimuli were presented on a Macintosh computer screen. At the start of each trial, a large fixation dot appeared in the middle of the screen for 1 s, after which time a word appeared in the center of the screen and remained there until the participant responded. Participants were instructed to be prepared to respond by resting one finger on the comma key and one on the adjacent period key, using their dominant hand. A "yes" response was made by pressing the comma key, and a "no" response by pressing the period key. After the participant responded, the word disappeared and the screen remained blank for 500 ms before the next trial began.

Participants were given three practice trials, followed by three blocks of experimental trials. The participant was instructed to make either a lexical decision or a semantic decision for each word and to respond as quickly and as accurately as possible. In the semantic-decision case, the participant was told to respond "yes" if the item represented something living, or part of a living thing, and "no" if otherwise. In the lexical-decision case, the participant was told to respond "yes" if the item was a word and "no" if otherwise.

Experimental trials were divided into three blocks, each consisting of a prime list and a target list. Each prime list within a block consisted of a randomized list of 25 semantic primes for the corresponding 5 primed targets in the following target list. Each target list in a block consisted of a randomized list of 15 words: 5 fillers, 5 unprimed targets, and 5 primed targets. This resulted in an average lag of 10 items between the last related prime word in a prime list and the corresponding target word in the following target list in the same block. There was a 2-min pause after each block of trials, as well as within each block between the prime and target lists. Which half of the target list was primed was counterbalanced across participants; thus, half of the participants' prime lists consisted of primes from Sublist A, and the other half's were from Sublist B. Participants were instructed to make the same decision for each word in each list. They were not told about the distinction between prime and target lists.

Results and Discussion

Trials on which an incorrect response was made were excluded from data analyses in this and all subsequent experiments. The mean reaction times (RTs) and error rates with standard errors are shown in the top half of Table 2. *T* tests on the error rates for each task revealed no significant differences. Thus, error rates were similar for primed and unprimed trials. The higher overall error rates on the semantic-decision task probably reflect the greater difficulty of this task, also reflected in the overall increase and greater variability in RTs. Because the distribution of participants' RT scores was skewed by some extremely long RTs, we analyzed the medians rather than the means in the following and all subsequent analyses. A three-way analysis of variance (ANOVA) with task (lexical or semantic decision) and sublist (A or B) as between-subjects variables and priming condition (primed or unprimed) as a within-subjects variable, conducted on the participants' median reaction times,

[4] One drawback to Seidenberg and McClelland's (1989) implementation of their model is that, unlike ours, it lacked a semantic layer. This may have contributed to its poor nonword reading performance. Subsequently, Plaut and McClelland (1993) have shown that attractor networks with a better orthographic representation are capable of reading nonwords with error rates and specific error patterns that are very similar to those made by humans.

Table 2
Mean Reaction Times (RTs; in Milliseconds) and Error Rates (in Percentages) for Experiments 1 and 2

Condition and measure	Task			
	Lexical decision		Semantic decision	
	M	*SE*	*M*	*SE*
Experiment 1				
Unrelated				
RT	597	18	759	23
Error	2	0.7	5	1.0
Related				
RT	587	18	716	24
Error	3	0.8	6	1.0
Experiment 2				
Unrelated				
RT			709	23
Error			6	1.1
Related				
RT			674	20
Error			6	1.2

showed no significant effect of word sublist, nor any significant interactions involving word sublist. The word sublist variable was therefore dropped from subsequent analyses. A 2 × 2 (Priming × Task) ANOVA on the same scores yielded significant main effects of priming, $F(1, 58) = 10.58$, $MSE = 21{,}533.8$, and of task, $F(1, 58) = 25.73$, $MSE = 634{,}889.3$), as well as a significant interaction between these variables, $F(1, 58) = 4.03$, $MSE = 8{,}208.8$. These priming effects are plotted in the right half of Figure 2. The interaction between priming and task type was investigated further by performing separate *t* tests on the data from the lexical-decision and semantic-decision tasks. There was a significant priming effect on the semantic-decision task, $t(29) = 2.97$, but not on the lexical-decision task, $t(29) = 1.30$, though there was a trend toward long-term priming on lexical decisions. In a 2 × 2 (Priming × Task) items ANOVA on the median scores for each word, treating items as participants, there were significant main effects of priming, $F(1, 29) = 11.72$, $MSE = 21{,}627.7$, and of task, $F(1, 29) = 90.33$, $MSE = 654{,}606.4$. Although the Task × Priming interaction was not significant in the items analysis, there was a trend in this direction, $F(1, 29) = 2.63$, $MSE = 11{,}213.3$, $p < 0.12$.

Because our target list included more inanimate than animate words, we also performed a 2 × 2 (Priming × Animacy) items ANOVA with animacy as a between-subjects variable on the semantic-decision scores. There was no significant effect of animacy, nor an Animacy × Priming interaction. Sign tests revealed that a significant proportion of the median scores were faster in the primed condition than in the unprimed condition for the animacy-decision task, for both participants (20/30) and words (23/30). For the lexical-decision task, the proportions of median scores showing priming did not differ from chance levels, for both participants (12/30) and words (17/30). In summary, as predicted, our semantic priming manipulation produced a long-term facilitatory effect spanning more than 2 min, evidenced predominantly on the semantic-decision task.

On average, participants' median semantic-decision times were 43 ms faster for primed words than for unprimed words (a 5.7% speedup), as compared with an insignificant 10-ms (1.7%) speedup on lexical decision. These results are qualitatively similar to the pattern of results seen in Simulation 1, where priming with blocks of words from two categories produced a 2.25% speedup on semantic settling time and an insignificant 0.88% speedup on orthographic settling time. Thus, we have demonstrated in both simulations and experiments that the long-term semantic priming effect depends critically on the degree of semantic involvement in the task.

One drawback to our experimental design is that in the lexical-decision condition, as in many lexical-decision studies, only the target lists contained nonwords; the priming lists contained none. Thus, it is possible that participants could have developed a bias toward making fast "yes" responses without fully processing the words. This could explain our failure to find long-term priming on the lexical-decision task. Joordens and Becker (1997) performed a similar lexical-decision experiment with much longer word lists of which 50% were nonwords. Primes and targets were separated by varying lags within blocks. The type of nonwords used for fillers was varied within subjects and between blocks to be more or less wordlike. In one condition, they were unpronounceable nonsense words—the least wordlike—in another, they were pronounceable nonsense words; and in the third, they were pseudohomophones (e.g., *brane*)—the most wordlike. As in our experiment, there was no semantic priming on lexical decision at lags greater than one for unpronounceable or pronounceable nonsense words. Of interest, in the pseudohomophone condition, there was long-lag priming, suggesting that when participants are forced to rely more on semantics, even lexical decisions can exhibit long-term semantic priming.

In Experiment 1, there were five primes per target. In our simulations, we demonstrated that the long-term semantic priming effect also holds when there is only a single prime per target, although there was a trend toward a smaller effect in Simulation 2. Experiment 2 was conducted to determine whether people behaved according to the model's predictions.

Experiment 2

In Experiment 2, we used the same materials as in Experiment 1 except that each participant was only presented with one of the five primes for each target word. It was predicted that there would still be long-term priming on semantic decisions, although the effect might be somewhat weaker. Because there was no evidence of semantic priming in lexical decision in Experiment 1, this time participants only performed the semantic-decision task.

Method

Participants. Thirty undergraduate students drawn from the same participant pool as for Experiment 1 participated in our study.

Materials. The same target words and prime words as in Experiment 1 were used, but each participant was presented with no more than one prime word per target. The 150 primes from Experiment 1 were divided into five sublists of 30 prime words, one for each of the 30 target words, and these sublists were further divided into primes corresponding to target Sublist A and Sublist B (Appendix D). Participants were each assigned to 1 of 10 corresponding groups (5 prime sublists × 2 target sublists), and 3 participants were run in each cell.

Procedure. The same procedure was used as in Experiment 1 for the semantic-decision condition except that this time each participant was presented with only one block of words, consisting of a list of 15 prime words followed by a list of 30 target words, with a 2-min pause in between. This resulted in an average lag of 21.5 items between a given prime in the prime word list and the corresponding word in the target list.

Results and Discussion

The mean RTs and percentage error rates with standard errors are shown in the bottom half of Table 2. These error rates are similar to those for the semantic-decision task in Experiment 1. A t test on these error rates revealed no significant difference between the primed and unprimed conditions. As before, semantic priming produced a facilitatory effect on the semantic-decision task. This priming effect is plotted in the right half of Figure 2. A two-way ANOVA on the participants' median scores with priming condition as a within-subjects variable and sublist (A,B) as a between-subjects variable revealed a significant main effect of priming, $F(1, 28) = 6.57$, $MSE = 18{,}200.4$. No other effects were significant. On average, participants were 35 ms faster in making semantic decisions for primed words than for unprimed words, a speedup of 4.9%. A one-way items ANOVA treating items as participants and priming condition as a within-subjects variable revealed a marginally significant priming effect, $F(1, 29) = 4.07$, $MSE = 11{,}956.8$. When animacy was added as a between-subjects variable, there were no significant effects. Sign tests revealed that a significant proportion of the median scores were faster in the primed than in the unprimed condition, for both participants (20/30) and words (22/30).

The priming effect on the semantic task in Experiment 2 was still surprisingly large at 4.9%, compared with 5.7% for Experiment 1. Because the two experiments used exactly the same methodology apart from the difference in the makeup of the priming blocks, we combined the semantic decision scores from Experiments 1 and 2 into a single two-way ANOVA, with priming as a within-subjects variable and experiment as a between-subjects variable. This analysis revealed once again a significant main effect of priming, $F(1, 58) = 15.66$, $MSE = 45{,}825.2$. Although there was a trend toward more priming in Experiment 1 (multiple primes per target) than in Experiment 2 (single prime per target), the interaction was not significant. These results parallel the pattern of results in Simulations 1 and 2.

The results of Experiments 1 and 2 are consistent with our interpretation of an automatic learning process underlying the long-term semantic priming effect. However, an alternative interpretation is that participants used recollective processes. Perhaps they explicitly recognized some targets as being related to previously seen items. Experiment 3 was designed to address this issue, as well as to investigate the decay of the effect.

Experiment 3

One way to improve our experimental procedure to minimize the possibility of strategic processing is to use much longer word lists. Experiment 3 was designed to determine whether the long-term semantic priming effect would hold up in longer word lists, while exploring the potential decay of the effect by systematically varying the prime–target lag. In this final experiment, we used a list of 48 prime–target word pairs and a slightly different procedure: prime and target words were interspersed within a single long block of trials. Three lags were examined: 0 items, 4 items, and 8 items. The ordering of stimuli was randomized, subject to two important constraints of (a) keeping response history constant across related and unrelated trials and (b) rotating the targets and primes through the two different relatedness conditions and the three different lags.

Method

Participants. Sixty undergraduate students drawn from the same participant pool as for Experiments 1 and 2 participated in our study.

Materials. Forty-eight prime–target pairs were chosen to maximize semantic relatedness. These items, as well as the practice and filler items, are listed in Appendix E. Some of the pairs were taken from the stimuli used in Experiment 2 of Joordens and Becker (1997), whereas others were provided by Timothy McNamara (personal communication, November 16, 1995). The filler words were selected to ensure that the proportions of animate and inanimate items were approximately equal. Each prime–target pair was coupled with another pair to allow the related prime for one target to serve as the unrelated prime for its coupled target. Four prime–target lists were created from these coupled prime–target pairs, so that each participant was exposed to only one of the four possible pairings of the four words. For example, *iceberg–glacier* was coupled with *fate–destiny*. Thus, a given participant would see either *iceberg* followed (after a lag of 0, 4, or 8 items) by *glacier*, *iceberg* followed by *destiny*, *fate* followed by *destiny*, or *fate* followed by *glacier*.

Procedure. In the present experiment, we used a running animacy-decision task in which participants make animacy decisions for a stream of stimuli, with each subsequent stimulus presented shortly after a decision is made for the current stimulus. Unlike the previous two experiments, there was no separation into a study phase and a test phase. Experimental trials were divided into eight blocks. Each block contained 12 words: 3 primes (which could be related or unrelated), 3 targets, and 6 filler words. It is important to note, however, that this block structure was not made apparent to the participants. From a participant's perspective, the experiment appeared to be structured as a single block of 116 trials. The first 20 of these trials were practice. The next 96 were experimental trials.

Table 3
Mean Reaction Times (RTs; in Milliseconds) and Error Rates (in Percentages) for Experiment 3

Condition and measure	Lag					
	0 Items		4 Items		8 Items	
	M	*SE*	*M*	*SE*	*M*	*SE*
Unrelated						
RT	868	36	859	23	883	32
Error	9	1.8	10	2.0	10	1.7
Related						
RT	776	18	817	23	812	30
Error	10	2.2	6	1.8	87	1.8

Given that we also wanted to rotate the stimulus pairs through the different lags, it was further required that there be three versions of each of the four lists described above: one in which the coupled prime–target pairs occurred at a lag of 0, one at a lag of 4, and one at a lag of 8. Thus, the complete counterbalancing of prime identity, target identity, and lag required that 12 separate lists be created. Response history across the related and unrelated versions of each lag was controlled within each of these lists in the following manner. First, four different blocks of 12 trials were created, each having a slightly different response history for the three lags. For example, the first block consisted of P8, F, P0, T0, F, F, P4, F, F, T8, F, and T4, where F is a filler word, P8 is the prime for Lag 8, T8 is the target for Lag 8, and so forth. Within each block, two of the primes were related and one was unrelated. Each block and its mirror version were presented within each list, with the mirror version of the block having identical response characteristics but a different "relatedness" associated with each lag. This ensured that every participant experienced one related and one unrelated trial at each of the four response histories within each of the three lags. Response history was not controlled across the different lags, given that the critical issue is to allow a controlled comparison of responses to the related and unrelated targets within each lag, not across the different lags.

The timing of events was as follows: First, the message *Press the ⟨space bar⟩ when you're ready to begin* was presented. Depression of the space-bar resulted in a 300-ms blank field followed by presentation of the first stimulus. Each response then initiated another 300-ms blank field followed by the subsequent stimulus, and so on, until all 116 stimuli had been presented and responded to.

Apparatus. Stimuli were displayed on a 15-inch (38.1-cm) SVGA color monitor and measured approximately 8 mm tall by 7 mm wide. Participants were positioned approximately 65 cm from the monitor and used the keyboard to enter their responses. Participants pressed the *1* key to indicate an "animate" decision and the *2* key to indicate an "inanimate" decision. They were instructed to respond as quickly and as accurately as possible.

Results and Discussion

Separate analyses were conducted on the mean RTs for correct word decisions and on the mean percent errors. Each analysis initially consisted of a 3 (lag) × 2 (context) ANOVA. *T* tests were then conducted to examine the reliability of the priming effect at each lag. The results of Experiment 3 are presented in Table 3 and Figure 4. With respect to the time it took participants to correctly categorize the targets as animate or inanimate, examination of Figure 4 suggests that priming occurred at all lags.

Statistical analysis of the RT data revealed only a main effect of context, $F(1, 59) = 18.30$, $MSE = 23{,}002.9$. Pairwise t tests revealed that priming effects were significant at all lags, all ts$(59) > 2.30$. An analogous ANOVA performed on the error data revealed a marginal effect of context, $F(1, 59) = 3.40$, $p < .08$. Planned t tests revealed that the only difference in error rate to reach significance across related and unrelated contexts was the 4% priming effect observed at Lag 4, $t(59) = 2.30$. It is important to note that this significant priming effect in errors at Lag 4 coincides with the smallest priming effect in RT.

The results of Experiment 3 extend those of Experiments 1 and 2. Whereas our first two experiments used relatively short word lists and allowed the prime–target lag to vary randomly, Experiment 3 used longer word lists and systematically varied the prime–target lag. The fact that long-term priming was still seen at the longest lag in our experiment supports our assumption that the basis of the priming effect is automatic learning rather than other strategic processes such as recollection. Further, our results indicate that this effect decays very little, if at all, over the range of lags we studied.

It should also be noted that there is considerable variability in RT scores for the animacy-decision task. Thus, the apparently smaller priming effect at Lag 4 compared with Lag 8 in the present experiment is likely to be due to chance or due to the fact that some portion of the priming effect manifested itself in errors for the Lag 4 condition as opposed to RT.

General Discussion

We began by making the novel prediction that it should be possible to produce long-term semantic priming. This prediction was generated by our computational model of priming and would not have been made without the model. Our experiments confirmed this prediction and are the first to demonstrate a semantic priming effect spanning many intervening items and lasting much longer than a few

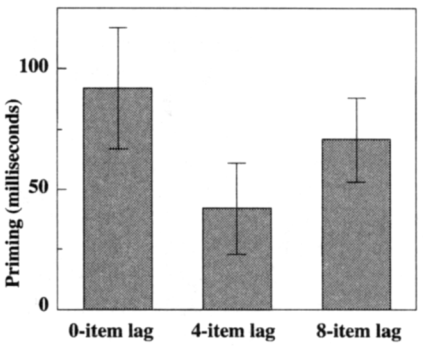

Figure 4. Plot of the unrelated minus the related reaction times (milliseconds) across prime–target lags for animacy decisions in Experiment 3. The bars represent the standard error of the priming effect.

seconds. Our theoretical account of long-term priming stated that priming involves changes in distributed representations that deepen the basins of attraction for primed words. Words that are semantically similar would also tend to have their attractors deepened because these attractors presumably overlap along many dimensions in semantic space. This led to the prediction that semantically similar words should benefit from priming on tasks involving semantic decisions because the system should be able to settle into the correct semantic attractor more quickly. However, because semantically similar words are usually unrelated orthographically, there should be less benefit on lower-level tasks like lexical decision. Results from both our simulations and experimental data support these predictions. To summarize, in Simulation 1 we showed significantly faster ST in the semantic but not in the orthographic layer in the network after priming with multiple semantically related primes. In Simulation 2 we showed the same pattern of results, when only one prime per target was used. Graphs of one-dimensional cross sections of the energy surface of the network revealed that our priming manipulation did indeed lower the basins of attraction for both primes and semantically related words. In Experiment 1 we replicated the results of Simulation 1 in human participants, who showed priming by multiple semantic primes on a semantic-decision task but not on a lexical-decision task. Note, however, that in both the simulations and the model there was a small trend toward long-term semantic priming on lexical decision. In Experiment 2, as predicted by the model, we found long-term semantic priming after only a single prime. Finally, in Experiment 3 we replicated the effect at lags of 0, 4, and 8 items by using a longer list of test items. Thus, our computational model and preliminary simulations were able to generate novel, testable predictions that guided our experimental work.

We would not necessarily expect the amount of priming to be in precise quantitative agreement between the simulations and the human experiments. Our simulations involved a small network of 269 units and a very small training set of only 40 words and 5 categories. In spite of these simplifications in the implemented version of our model, the main results are in surprisingly close qualitative agreement with the experimental data. Thus, we have suggestive results from our simulations and converging evidence from experiments that are consistent with our prediction that long-term priming involves changes in connection strengths in distributed representations that alter the attractor basins of primes and related words.

Our model suggests one explanation as to why previous experiments in the literature failed to produce long-term semantic priming on lexical decision: the match between the representational level of prime–target overlap (e.g., orthographic vs semantic) and the level of processing in the test task is a critical variable. If the prime and target words have overlapping basins of attraction only at the semantic level, for example, we would expect to see evidence of priming only on tasks that engage semantic processing to a sufficient degree. Lexical decisions on words embedded in a background of scrambled nonwords are easy to perform quickly without doing any semantic retrieval. However, as Joordens and Becker (1997) have shown, when the lexical-decision task is made more difficult by making the nonwords more wordlike, semantics does appear to be important for making accurate decisions, and thus long-term semantic priming is observed even on this task. Note that our experiments do not allow us to determine whether it is making animacy decisions on the prime, the target, or both, that is necessary to observe long-term semantic priming. Our model predicts that both are critical, as is consistent with the literature on long-term repetition priming (e.g., Vriezen et al., 1995).

Other Forms of Long-Term Priming

We now examine two other forms of long-term priming, morphological and conceptual priming, both of which can be accommodated within the framework of our model. We then compare the major theoretical accounts of long-term priming with our own.

Morphological priming. Morphological priming in lexical decision has been reported at long lags in a variety of languages including English (Stolz & Feldman, 1995), Hebrew (Bentin & Feldman, 1990), Serbo-Croatian (Feldman & Moskovljević, 1987) and American sign language (Emmorey, 1991). These findings might seem to be at odds with our results. Why should it be so easy to produce long-term priming for morphologically but not for semantically related words on this task? To address this question, one needs to consider the levels of representation of semantic and morphological relatedness. Many models of lexical access (see, e.g., Chialant & Caramazza, 1995) assume that words are decomposed into morphologically based constituents at a presemantic, lexical level of representation. If our network model was trained on a much larger word corpus, it might be expected to discover some aspects of morphological structure as a useful intermediate-level representation in mapping from orthography to semantics and phonology. Thus, the hidden layer subsequent to the orthographic layer in our network could serve this role, among other things. It should be noted, however, that whether there is in fact a separate level of morphological representation is a matter of debate in the literature (see P. T. Smith, 1995, for example, for a dissenting opinion and some simulation results).

If the above view of morphological representation is correct, and if we are correct in assuming that lexical decisions are based primarily on the ST of presemantic units, then the lexical-decision task would depend directly on the ST of morphological representational units but would not depend directly on semantic-level activation. This would explain why long-term priming is seen in lexical decisions for morphologically similar words but is not normally seen for semantically similar words. However, this argument does not preclude the possibility that the lexical-decision task is sensitive to semantic influences.[5] In fact, there is plenty of

[5] Further, it is important to note that morphologically related words are usually also semantically related. Hence, in Rueckl, Mikolinski, Raveh, Miner, and Mars's (1997) model of morphological priming, the semantic to orthographic pathway plays an important role.

experimental evidence to support this possibility (reviewed in Balota, 1990). Our network model explicitly incorporates semantic influences on lexical decision through feedback connections from the semantic to the orthographic layer via a hidden layer. Thus, Joordens and Becker (1997) found that under conditions thought to increase the importance of semantics in making lexical decisions, long-term semantic priming can be obtained.

Conceptual priming. Findings in the literature of long-term conceptual priming (reviewed in Roediger, 1990; Roediger & McDermott, 1993) are consistent with our theoretical account of long-term semantic priming. These studies have found long-term priming on conceptually driven tasks such as category exemplar generation (e.g., the word *dog* is presented during a study or priming phase and is later generated in response to the cue *animal*). An attractor network could account for such an effect in the following way: The system settles into an attractor state in response to the prime word, and long-term priming results. This deepens the basin of attraction for the prime, which overlaps with that of the cue, making the same response more probable or more rapidly generated when the cue is subsequently encountered. Thus, we view long-term semantic priming and conceptual priming as having very similar underlying mechanisms. Future developments of the model will address how long such effects can last and how they can be reinstated in the appropriate context.

Woltz (1990, 1996) has studied a more complex form of conceptual–repetition priming in which the participants' task is to judge whether pairs of words (e.g., *moist–damp*) have the same meanings. When the same pair of words reappeared later in the list, substantial long-lag priming was observed. Moreover, when the prime pair preceded a synonymous target pair (e.g., *moist–damp* followed by *soggy–wet*), there was also long-lag priming. This finding could be accommodated within our framework for long-term semantic priming, assuming each word's attractor basin in the prime pair was individually deepened, allowing faster processing of the target pair. However, results in other conditions of Woltz's (1990) second experiment do not fit this explanation. Specifically, when either the prime or target pair was unrelated, even if corresponding elements of the prime and target pairs were synonymous (e.g., *ample–enclose* followed by *enough–surround*), there was no evidence of long-lag priming. Further, Woltz's (1996) second experiment included some identical pairs of primes and targets that were synonymous (e.g., *moist–moist* followed by *damp–damp*), with sequential word presentation to ensure full semantic processing of at least the first member of each pair; in this condition, there was no long-lag priming. Woltz concluded from these data that priming on his semantic-comparison task primarily involves the particular processes involved in comparing word meanings, rather than a more general reinforcement of semantic-level representations. Although these data appear to fly in the face of our semantic learning explanation, it may be that the added complexity of Woltz's task relies heavily on postsemantic-retrieval processing and therefore is insensitive to semantic learning effects. More specifically, the bottleneck in this task may be the word comparison stage, in which case any benefits of enhancing semantic-level representations would be small in comparison to effects of enhancing the particular connections involved in making the same–different judgment. Thus, we would predict that if participants engaged in semantic comparisons on the prime list, followed by simple semantic decisions on the target list (e.g., a same–different meaning decision on *moist–moist* followed by an animacy decision on *damp*), one would indeed see long-lag priming. We now consider how our model could be extended to allow such processing-specific learning effects as reported by Woltz.

Transfer-appropriate processing accounts. Two important theoretical accounts of repetition priming in the literature deal directly with task-specificity effects during the priming–study phase: Tulving and Schacter's (1990) multiple memory systems theory postulates that different brain subsystems are involved in different priming tasks, whereas Roediger's (1990) transfer appropriate processing theory postulates that different kinds of processing are required for different tasks. Both accounts predict that a critical factor in producing priming is a match between the level or type of processing involved when the participant is responding to primes versus targets. In particular, they predict task-specific priming effects for data-driven versus conceptually driven tasks. Vriezen et al. (1995) observed task-specific repetition priming effects even within the semantic domain. Repetition priming was observed when the task was a size decision (*is it bigger than a bread box?*) at study and a dimension decision (*is it wider than it is long?*) at test or vice versa. But there was no repetition priming if the judgments at study and test involved different semantic domains, that is, size versus animacy.

We are sympathetic to the transfer-appropriate processing accounts described above and view them as being compatible with our explanation of long-term priming. Given that participants apparently can truncate their level of processing as necessary according to task demands, one would expect the level of long-term priming effects to be in correspondence with the level of processing. To account for within-level processing-specificity effects such as those found by Vriezen et al. (1995) and Woltz (1990, 1996), we would have to extend our model to include some sort of selection procedure to focus processing on a particular layer or subset of units during priming. One way to to do this would be to add a response layer to the network and explicitly train it to respond according to the current task demands. It might learn to make animacy decisions, for example, by inhibiting processing of semantic features irrelevant to this task. The network would then be prevented from strengthening the connections between features relevant to the animacy decision and other irrelevant features. Hence, the network would exhibit no transfer from the animacy decision to unrelated tasks such as size decisions. A number of studies by Dagenbach, Carr, and colleagues (reviewed in Carr et al., 1994, pp. 704–705) provided evidence in support of such an attentional mechanism, at least during memory retrieval. They postulated that a center-surround attentional mecha-

nism modulates the retrieval of weakly activated items, thus inhibiting irrelevant associated items.

Recollective processes. An alternative explanation of our findings is that the long-term effects we found are due to learning in an episodic memory system or to some sort of strategic processing. For example, perhaps participants noticed the semantic relationship between the target words and previously encountered words, and this awareness facilitated their semantic decision making. This is in contrast to our view that long-term priming involves incremental learning in a distributed memory system and that the transfer of priming effects to similar words is due to automatic generalization of this learning to words with overlapping basins of attraction. Such a mechanism should not require conscious strategies, although they might play an additional role. Further, it may turn out that conscious processing of the prime is necessary for long-term learning—even without the necessity of laying down an episodic memory trace.

Ratcliff and McKoon (1988) have questioned the notion that all long-term priming effects are subserved by the same mechanism. For example, Ratcliff et al. (1985) have reported an intermediate-range priming effect for newly learned associates in recognition memory tasks whose time course appears to differ from both short-term and long-term priming. Whereas in lexical decision, semantic–associative priming is typically found to decay rapidly after one or two intervening items, in recognition priming, the decay (for repeated items and for newly learned associates as primes) is much more gradual but is largely diminished by a lag of 8 items. It is possible that this effect is due to the use of conscious strategies. Our procedure was designed to minimize the use of conscious strategies by forcing participants to respond quickly to serially presented items. However, it is possible that our word lists (Experiments 1 and 2) or lags (Experiment 3) were short enough that participants did in fact use recollective strategies.

Short-Term Semantic Priming

Having discussed various findings and theoretical accounts concerning long-term priming, we now turn to the literature on short-term semantic priming. Another possible interpretation of our findings is that they are qualitatively the same as traditional short-term effects. Perhaps existing theories of short-term priming could be extended to accommodate our findings. But before turning to these theoretical accounts, we discuss another important distinction in the semantic priming literature: between semantic and associative relatedness. Most theories make different predictions depending on whether words are purely semantically related (e.g., *bread–cake*) or are both semantically and associatively related (e.g., *bread–butter*).

Semantic versus associative priming. Whether short-term semantic priming occurs at all for words that are purely semantically related but not associatively related has been a subject of debate in the literature. Although some studies have reported pure semantic priming on lexical decision (e.g., Fischler, 1977; Lupker, 1984), Shelton and Martin (1992) found no such effect on this task with a procedure thought to minimize strategic processing.[6] However, although Moss, Ostrin, Tyler, and Marslen-Wilson (1995), like Shelton and Martin, did not find pure semantic priming on visual lexical decision for category coordinates (e.g., *pig–horse*), they did find pure semantic priming for instrument relations (e.g., *broom–floor*). Further, they observed pure semantic priming for all types of semantic relations tested on auditory lexical decision. In a more recent unpublished experiment, R. Martin, J. Altarriba, S. Wayland, and J. Shelton (personal communication, December 21, 1995) attempted to replicate the Shelton and Martin (1992) study by testing students at Rice University (the site of the original study) and at other universities. Although they again failed to obtain priming on the semantically related pairs with Rice students as participants, they did obtain significant priming for students at the other universities with the same materials and procedures. These investigators noted that the students at the other universities had substantially longer RTs, which may have related to the differing results. One account of these findings is that the semantic level of representation takes longer to activate than the orthographic–lexical level because it is a later stage of processing. Thus, slower responding participants would be activating semantics to a larger degree and would therefore be more likely to show semantic priming on lexical decision.

In a recent set of studies, McRae and Boisvert (1997) used word pairs that had been assessed for semantic overlap on the basis of empirically derived feature norms (McRae, de Sa, & Seidenberg, 1997) and which were not associatively related according to word-association norms. They found significant short-term semantic priming for both semantic decisions (*does the word refer to something you can touch?*) and lexical decisions. They also tested the semantically related word pairs used in Shelton and Martin's (1992) studies and replicated their null result. In fact, according to McRae and Boisvert's similarity rating norms, these words were not highly semantically related. Taken together, these studies clearly demonstrate the importance of a strong overlap in semantics for obtaining short-term "pure semantic" priming. These data support models of word recognition that use distributed representations of semantics, like the Hinton–Shallice (1991) model on which our account of long-term priming is based. As we have argued, sufficient semantic overlap should also be a critical feature for long-term semantic priming. In fact, we would predict long-term priming if the highly related prime–target pairs in McRae and Boisvert's study were presented at long lags. This should be particularly evident on a semantic-decision task.

The above data actually raise a tricky issue for our account of the short- versus long-term distinction in semantic priming: We have argued that previous attempts to get long-term semantic priming in lexical decision failed because the task is not sufficiently semantic in nature. If this is true, then why should there be greater short-term priming in lexical decision as the semantic overlap between primes and

[6] Plaut (1995) has proposed a connectionist model that accounts for Shelton and Martin's (1992) data.

targets is increased and as participants' RTs become longer? Further, studies that directly manipulated the depth of processing of the prime have shown that the semantic priming effect is attenuated (Kaye & Brown, 1985) or eliminated (M. C. Smith, Theodor, & Franklin, 1983; Henik, Friedrich, & Kellog, 1983; Friedrich, Henik, & Tzelgov, 1991) when the prime is processed in a letter search task. Presumably, in such shallow processing tasks semantic processing is virtually absent. Taken together, these data suggest that some amount of semantic processing is a necessary condition for short-term pure semantic priming. One possible resolution of these data with our account is to assume that in order for long-term priming to occur, the relevant semantic-level nodes must be fully activated or activated above some threshold. An explicitly semantic task such as animacy decision would achieve this, whereas a lexical-decision or naming task may or may not. Evidence in support of the assumption that sufficiently strong activation is required for long-term priming comes from studies showing that subliminal prime presentation eliminates long-term repetition priming in lexical decision (Forster & Davis, 1984), stem completion, and fragment completion (Forster, Booker, Schacter, & Davis, 1990). Further, as discussed following our first experiment, Joordens and Becker (1997) indirectly manipulated the depth of processing of primes and targets in a running lexical-decision task by varying the word–nonword similarity and found long-term priming only for deeply processed primes and targets.

Contextual priming. Recent studies by Hess, Foss, and Carroll (1995) suggested that semantic–associative relatedness may be an overly simplistic explanation of the richness of contextual priming effects observed in more realistic reading situations. Their findings indicate that when a word is read in the context of a related sentence, this broader context predicts priming more consistently than does the single preceding word. These findings have clear implications for both short-term priming studies and studies such as ours involving long-term priming in isolated word recognition. Clearly, when one is reading a passage of text as compared with a list of isolated words, a far richer base of information is available to facilitate word processing. These facilitatory effects may indeed turn out to be long-term as well as short-term. We now turn to two major theoretical accounts of short-term priming: residual activation theories and compound-cue theory.

Activation Theories

Many theories have been proposed to account for short-term semantic–associative priming that involve some sort of residual activation effect from the prime, for example, by the activation of an "expectancy set" of word nodes representing possible associates of the prime (e.g., C. Becker, 1980), automatic spreading activation in a localist network (Anderson, 1983; Collins & Loftus, 1975; Collins & Quillian, 1969; Neely, 1977; Posner & Snyder, 1975), or sustained activation of a distributed representation (A. Sharkey & N. Sharkey, 1992; N. Sharkey, 1989; Masson, 1989, 1991, 1995). See Neely (1991) for an excellent survey of the major findings and theoretical accounts in the enormous literature on short-term semantic–associative priming. We focus on the localist and distributed activation accounts here because they have two things in common with our account of priming: They specifically address the time course of semantic priming, and they are based on similar models of word recognition that involve activation patterns in networks of interconnected nodes.

As mentioned in the introduction, the standard finding is that semantic priming fully dissipates for prime–target lags greater than one (see, e.g., Joordens & Besner, 1992). Spreading activation accounts of priming (Anderson, 1983; Collins & Loftus, 1975; Collins & Quillian, 1969; Neely, 1977; Posner & Snyder, 1975) assume a localist representation in which there is a node for each word or concept. When a word's node is activated, activity automatically spreads across links to related word nodes, with activation decaying with distance and time. Typically, relatedness is defined in terms of free association norms (see, e.g., McNamara, 1992a, 1992b). Priming occurs when the residual activation of a target node reduces the time required to fully activate it by the target word. Priming across a lag of 1 can be accommodated easily in these models because of the localist representation. Provided the intervening word is unrelated, it should not interfere with activation of the prime or target word node. Spreading activation models also predict mediated priming (McNamara, 1992a), in which concepts related to the prime via two-step and three-step chains would be weakly activated and therefore weakly primed. However, Ratcliff and McKoon (1994) noted that each word has many associates; therefore, they argued that if each word activates about 20 related nodes, in a three-step chain, 8,000 nodes would be activated, constituting a substantial proportion of the adult lexicon. Spreading activation models could account for long-lag priming as follows: If some activation spreads from a prime word to its semantically related target and that activation decays very slowly, it could persist across arbitrarily many intervening items. However, by Ratcliff and McKoon's line of reasoning, if substantial activation were to persist beyond two intervening items, and each active node in turn activated its own 20 or so associates, one would soon have the entire lexicon activated.

Distributed connectionist models provide a somewhat similar account of short-term priming in terms of residual activation, but by using a different representation of semantics. Instead, each word or concept is represented by a pattern of activation across a large set of nodes such that related concepts activate overlapping representations. Another fundamental difference is that typically only one concept can be fully activated at one time, although multiple concepts could be weakly activated. For the sake of brevity, we focus on Masson's (1995) model, noting that Sharkey and Sharkey (A. Sharkey & N. Sharkey, 1992; N. Sharkey, 1989) proposed a very similar model. Masson accounted for short-term semantic priming by using a connectionist model with an input layer representing orthographic features of a word, a meaning layer representing the word's semantics, and a phonological layer representing the word's pronunciation. When a word is processed by the network, residual

activation remains at the semantic layer. Thus, when a new pattern is presented to the input units, the initial state of the semantic units can influence the time the network takes to settle to a stable response. If the initial state is similar to the network's final response, ST should be fast relative to starting from an unrelated semantic state.

Masson's (1995) model has greater difficulty than a localist one in handling priming across a nonzero lag. Activation of an intervening item's distributed representation would presumably wash out that of the prime. However, Masson's simulations show that if the intervening item only partially activates the semantic layer, the residual activation is noisier but still sufficient to produce priming. Further, Masson's model handles the differential effects of interposing an unrelated versus a neutral prime. When an unrelated prime is processed, the network is far from the correct state for the target, and therefore retrieval is slow. In contrast, a neutral prime such as a row of *X*s would not activate a systematic pattern on the semantic layer and thus would not interfere with semantic priming.

Masson's (1995) account of semantic priming is based on a model of word recognition very similar to our model. However, both Masson's and the spreading-activation accounts of the mechanism of priming differ substantially from ours. Any sustained activation account must predict that some decay occurs. As new stimuli are encountered, their activations must predominate over the residual effects of previous stimuli. Thus, such effects can only be very short-lived and ultimately must decay down to zero. Although the exact time course of decay of activation is a subject of ongoing empirical investigation, it is generally agreed that this decay is very rapid and could not possibly accommodate the long-lasting semantic priming effects obtained in our studies.

If models of short-term semantic priming cannot reasonably accommodate very long-lasting effects, perhaps our model of long-term priming could be made to accommodate short-term effects. For example, perhaps the plastic changes associated with long-term priming are initially very large but then decay down to some stable baseline, as proposed by McClelland and Rumelhart (1986). This view has a certain appeal, to the extent that it is parsimonious. However, our experience with training neural networks with very large weight changes suggests that this is a recipe for disaster. That is, large changes in weights to accommodate one particular stimulus tend to produce massive interference effects with subsequent stimuli, along with very high error rates. This pattern of results is not typical of human participants in priming experiments. Thus, unless the learning is mediated by a specialized memory system that is relatively immune to interference (e.g., the episodic system) it is likely to involve fairly small, incremental adjustments (McClelland, McNaughton, & O'Reilly, 1995; O'Reilly & McClelland, 1994). Our conclusion is that a combination of short-term activation effects and long-term weight changes seems the most plausible account of the semantic priming literature.

Compound-Cue Theory

In contrast to explicit activation accounts of short-term priming, compound-cue accounts (e.g., Dosher & Rosedale, 1989; Ratcliff & McKoon, 1988; Whittlesea & Jacoby, 1990) view priming as a result of the combination of the prime and target into a composite cue in short-term memory. This compound cue produces greater familiarity than either cue alone and hence a speeded response. Depending on what sort of processing participants engaged in between successive items, such a mechanism could conceivably produce priming across arbitrarily long lags, though for very long stimulus lists the effect would presumably drop off sharply as the lag increased.

However, when one considers how such a mechanism might actually be implemented within a distributed network model of memory, it begins to look very much like residual activation accounts. For example, Ratcliff and McKoon (1994) have proposed that in a network such as the Seidenberg and McClelland (1989) model of lexical decision,

> . . . gradual (stochastic) replacement of one item by the next item . . . would allow the representation at input to be a compound . . . [which] could percolate through the whole network. To produce semantic priming effects, . . . the semantic layer could represent semantic feature overlap, so that a compound of related items would produce a better match to memory and faster responses. (p. 183)

It is difficult to distinguish this proposal from that of Masson (Masson, 1991, 1995) or Sharkey (A. Sharkey & N. Sharkey, 1992; N. Sharkey, 1989). Like the Masson and Sharkey and Sharkey models, we consider compound-cue accounts to be a viable explanation for short time-scale priming effects, complementary to our account of long time-scale effects.

Conclusions

Our view of long-term priming as incremental learning differs from previous accounts of short-term semantic–associative priming in the literature that involve transient effects like residual activation or expectancy generation. These accounts were developed to explain short-lived effects like priming by associatively related words (e.g., *bread* and *butter*).[7] Such processes may help the cognitive system to process language rapidly by preparing it for what is likely to occur next, on the basis of the current context (but see Hess et al., 1995). Long-term learning, in contrast, is presumably necessary to lay down cumulative memory traces of what a person has experienced over a lifetime. This view suggests that long-term priming reflects the normal course of learning. Presumably these different short- and

[7] It is widely believed that such effects reflect priming at the lexical level rather than at the semantic level. For example, de Groot (1990) found no differences in short-term priming for associatively related words in the lexical-decision and animacy-decision tasks, except in one experiment in one condition. In that condition, the effect only showed up for animate words, and subsequent experiments implicated strategic processing as a probable explanation of this result.

long-term mechanisms would act in concert in a complete cognitive system.

The data reported here highlight the value of combining modeling and experimental approaches to developing theoretical accounts of cognitive processes. Having a computational model allows us to postulate in a precise and specific way the mechanisms that may underlie long-term priming. The process of building our priming model led us to make novel predictions about long-term semantic priming, which we were then able to confirm experimentally. In our simulations, we have undoubtedly used a highly oversimplified representation of human semantic memory, as well as of the semantic learning process. Thus, we could not hope for exact quantitative agreement with human data. However, for the most part we were able to obtain good agreement between our qualitative pattern of simulation results and the pattern of priming in human participants. From this we can conclude that the model did capture the essential aspects of orthographic and semantic memory organization that are relevant to the lexical- and semantic-decision tasks. Thus, the general approach appears to be promising as a way of bridging the gap between neural modeling and experimental approaches to understanding human cognition.

References

Anderson, J. R. (1983). *The architecture of cognition.* Cambridge, MA: Harvard University Press.

Balota, D. A. (1990). The role of meaning in word recognition. In D. A. Balota, G. B. F. d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 9–32). Hillsdale, NJ: Erlbaum.

Balota, D. A., & Paul, S. (1996). Summation of activation: Evidence from multiple primes that converge and diverge within semantic memory. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 827–845.

Battig, W., & Montague, W. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology Monograph, 80* (3, Pt. 2).

Bavelier, D., & Jordan, M. I. (1993). A dynamic model of priming and repetition blindness. In S. J. Hanson, J. D. Cowan, & C. L. Giles (Eds.), *Advances in neural information processing systems 5* (pp. 879–886). San Mateo, CA: Morgan Kaufmann.

Becker, C. (1980). Semantic context effects in visual word recognition: An analysis of semantic strategies. *Memory & Cognition, 8,* 493–512.

Becker, S., Behrmann, M., & Moscovitch, M. (1993). Word priming in attractor networks. In *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society* (pp. 231–236). Hillsdale, NJ: Erlbaum.

Bentin, S., & Feldman, L. (1990). The contribution of morphologic and semantic relatedness to repetition priming at short and long lags: Evidence from Hebrew. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 42*(A), 693–711.

Bentin, S., & Moscovitch, M. (1988). The time course of repetition effects for words and unfamiliar faces. *Journal of Experimental Psychology: General, 117,* 148–160.

Besner, D., Twilley, L., McCann, R. S., & Seergobin, K. (1990). On the association between connectionism and data: Are a few words necessary? *Psychological Review, 97,* 432–446.

Bower, G. (1996). Reactivating a reactivation theory of implicit memory. *Consciousness and Cognition, 5,* 27–72.

Carr, T., Dagenbach, D., VanWieren, D., Radvansky, L., Alejano, A., & Brown, J. (1994). Acquiring general knowledge from specific episodes of experience. In C. Umiltá & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and nonconscious information processing* (pp. 697–724). Cambridge, MA: MIT Press.

Chialant, D., & Caramazza, A. (1995). Where is morphology and how is it processed? The case of written word recognition. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 55–76). Hillsdale, NJ: Erlbaum.

Collins, A. M., & Loftus, E. F. (1975). A spreading activation theory of semantic processing. *Psychological Review, 82,* 407–428.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior, 8,* 240–247.

Dannenbring, G., & Briand, K. (1982). Semantic priming, and the word repetition effect in lexical decision. *Canadian Journal of Psychology, 36*(3), 435–555.

de Groot, A. M. B. (1990). The locus of the associative-priming effect in the mental lexicon. In D. A. Balota, G. B. F. d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 101–123). Hillsdale, NJ: Erlbaum.

Dosher, B. A., & Rosedale, G. (1989). Integrated retrieval cues as a mechanism for priming in retrieval from memory. *Journal of Experimental Psychology: General, 118,* 191–211.

Emmorey, K. (1991). Repetition priming with aspect agreement morphology in American sign language. *Journal of Psycholinguistic Research, 20,* 365–388.

Feldman, L. B., & Moskovljević, J. (1987). Repetition priming is not purely episodic in origin. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 573–581.

Feustel, T. C., Shiffrin, R. M., & Salasoo, A. (1983). Episodic and lexical contributions to the repetition effect in word identification. *Journal of Experimental Psychology: General, 112,* 309–346.

Fischler, I. (1977). Semantic facilitation without association in a lexical decision task. *Memory & Cognition, 5,* 335–339.

Forster, K. I. (1987). Form-priming with masked primes: The best-match hypothesis. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 127–146). Hillsdale, NJ: Erlbaum.

Forster, K. I., Booker, J., Schacter, D. L., & Davis, C. (1990). Masked repetition priming: Lexical activation or novel memory trace? *Bulletin of the Psychonomic Society, 28,* 341–345.

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 680–698.

Friedrich, F. J., Henik, A., & Tzelgov, J. (1991). Automatic processes in lexical access and spreading activation. *Journal of Experimental Psychology: Human Perception and Performance, 17,* 792–806.

Graf, P., & Schacter, D. L. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11,* 501–518.

Hebb, D. O. (1949). *The organization of behavior.* New York: Wiley.

Henderson, L., Wallis, J., & Knight, D. (1984). Morphemic structure and lexical access. In H. Bouma & D. Bouwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 211–224). Hillsdale, NJ: Erlbaum.

Henik, A., Friedrich, F. J., & Kellog, W. A. (1983). The dependence of semantic relatedness effects upon prime processing. *Memory & Cognition, 11,* 366–373.

Hess, D. J., Foss, D. J., & Carroll, P. (1995). Effects of global and local context on lexical processing during language comprehension. *Journal of Experimental Psychology: General, 124,* 62–82.

Hinton, G. E. (1989). Deterministic Boltzmann learning performs steepest descent in weight-space. *Neural Computation, 1,* 143–150.

Hinton, G. E., & Shallice, T. (1991). Lesioning an attractor network: Investigations of acquired dyslexia. *Psychological Review, 98,* 74–95.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences U.S.A., 79,* 2554–2558.

Hopfield, J. J. (1984). Neurons with graded response have collective computational properties like those of two-state neurons. *Proceedings of the National Academy of Sciences U.S.A., 81,* 3088–3092.

Jacoby, L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General, 110,* 306–340.

Joordens, S., & Becker, S. (1997). The long and short of semantic priming effects in lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 1083–1105.

Joordens, S., & Besner, D. (1992). Priming effects that span an intervening unrelated word: Implications for models of memory representation and retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 483–491.

Kaye, D. B., & Brown, S. W. (1985). Levels and speed of processing effects on word analysis. *Memory & Cognition, 13,* 425–434.

Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. P. (1983, May 13). Optimization by simulated annealing. *Science, 220,* 671–680.

Lupker, S. (1984). Semantic priming without association: A second look. *Journal of Verbal Learning and Verbal Behavior, 23,* 709–733.

Masson, M. E. J. (1989). Lexical ambiguity resolution in a constraint satisfaction network. In *Proceedings of the Eleventh Annual Conference of the Cognitive Science Society* (pp. 757–764). Hillsdale, NJ: Erlbaum.

Masson, M. E. J. (1991). A distributed memory model of context effects in word identification. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 233–263). Hillsdale, NJ: Erlbaum.

Masson, M. E. J. (1995). A distributed memory model of semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 3–23.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review, 102,* 419–457.

McClelland, J. L., & Rumelhart, D. E. (1986). A distributed model of human learning and memory. In J. L. McClelland, D. E. Rumelhart, & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 2, pp. 170–215). Cambridge, MA: Bradford Books.

McNamara, T. P. (1992a). Priming and constraints it places on theories of memory and retrieval. *Psychological Review, 99,* 650–662.

McNamara, T. P. (1992b). Theories of priming: I. Associative distance and lag. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 1173–1190.

McRae, K., & Boisvert, S. R. (1997). *Automatic semantic similarity priming.* Manuscript submitted for publication.

McRae, K., de Sa, V., & Seidenberg, M. S. (1997). On the nature and scope of featural representations in word meaning. *Journal of Experimental Psychology: General, 126,* 99–130.

Meyer, D. M., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90,* 227–234.

Meyer, D. M., Schvaneveldt, R. W., & Ruddy, M. G. (1974). Functions of graphemic and phonemic codes in visual word-recognition. *Memory & Cognition, 2,* 309–321.

Monsell, S. (1985). Repetition and the lexicon. In A. W. Ellis (Ed.), *Progress in the psychology of language* (pp. 147–195). London: Erlbaum.

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review, 76,* 165–178.

Moss, H. E., Ostrin, R. K., Tyler, L. K., & Marslen-Wilson, W. D. (1995). Accessing different types of lexical semantic information: Evidence from priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 863–883.

Neely, J. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General, 106,* 226–254.

Neely, J. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264–336). Hillsdale, NJ: Erlbaum.

O'Reilly, R., & McClelland, J. (1994). Hippocampal conjunctive encoding, storage, and recall: Avoiding a tradeoff. *Hippocampus, 4,* 661–682.

Peterson, C., & Anderson, J. R. (1987). A mean field theory learning algorithm for neural networks. *Complex Systems, 1,* 995–1019.

Plaut, D. C. (1995). Semantic and associative priming in a distributed attractor network. In *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society* (pp. 37–42). Hillsdale, NJ: Erlbaum.

Plaut, D. C., & McClelland, J. L. (1993). Generalization with componential attractors: Word and nonword reading in an attractor network. In *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society* (pp. 824–829). Hillsdale, NJ: Erlbaum.

Plaut, D. C., & Shallice, T. (1993a). Deep dyslexia: A case study of connectionist neuropsychology. *Cognitive Neuropsychology, 10,* 377–500.

Plaut, D. C., & Shallice, T. (1993b). Perseverative and semantic influences on visual object naming errors in optic aphasia: A connectionist account. *Journal of Cognitive Neuroscience, 5,* 89–117.

Posner, M. I., & Snyder, C. R. R. (1975). Facilitation and inhibition in the processing of signals. In P. M. Rabbit (Ed.), *Attention and performance V* (pp. 669–681). New York: Academic Press.

Ratcliff, R., Hockley, W., & McKoon, G. (1985). Components of activation: Repetition and priming effects in lexical decision and recognition. *Journal of Experimental Psychology: General, 114,* 435–450.

Ratcliff, R., & McKoon, G. (1988). A retrieval theory of priming in memory. *Psychological Review, 95,* 385–408.

Ratcliff, R., & McKoon, G. (1994). Retrieving information from memory: Spreading-activation theories versus compound-cue theories. *Psychological Review, 101,* 177–184.

Richardson-Klavehn, A., & Bjork, R. A. (1988). Measures of memory. *Annual Review of Psychology, 39,* 475–543.

Roediger, H. L., III, (1990). Implicit memory: Retention without remembering. *American Psychologist, 45,* 1043–1056.

Roediger, H. L., III, & McDermott, K. B. (1993). Implicit memory in normal human subjects. In H. Spinnler & F. Boller (Eds.), *Handbook of neuropsychology* (Vol. 8, pp. 63–131). Amsterdam: Elsevier.

Rueckl, J. G. (1990). Similarity effects in word and pseudoword repetition priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 374–391.

Rueckl, J. G. (1995). Ambiguity and connectionist networks: Still settling into a solution—Comment on Joordens and Besner (1994). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 501–508.

Rueckl, J. G., Mikolinski, M., Raveh, M., Miner, C. S., & Mars, F. (1997). Morphological priming, fragment completion and connectionist networks. *Journal of Memory and Language, 36,* 382–405.

Rueckl, J. G., & Olds, E. M. (1993). When pseudowords acquire meaning: The effect of semantic associations on pseudoword repetition priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 3,* 515–527.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986, October 9). Learning internal representations by back-propagating errors. *Nature, 323,* 533–536.

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 501–518.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed developmental model of word recognition and naming. *Psychological Review, 96,* 523–568.

Sharkey, A., & Sharkey, N. (1992). Weak contextual constraints in text and word priming. *Journal of Memory and Language, 31,* 543–572.

Sharkey, N. (1989). The lexical distance model and word priming. In *Proceedings of the Eleventh Annual Conference of the Cognitive Science Society* (pp. 860–867). Hillsdale, NJ: Erlbaum.

Shelton, J., & Martin, R. (1992). How semantic is automatic semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 1191–1210.

Sloman, S., Hayman, C., Ohta, N., Law, J., & Tulving, E. (1988). Forgetting in primed target completion. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 223–239.

Smith, M. C., Theodor, L., & Franklin, P. E. (1983). The relationship between contextual facilitation and depth of processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 697–712.

Smith, P. T. (1995). Are morphemes really necessary? In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 365–382). Hillsdale, NJ: Erlbaum.

Stolz, J. A., & Feldman, L. B. (1995). The role of orthographic and semantic transparency of the base morpheme in morphological processing. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 109–129). Hillsdale, NJ: Erlbaum.

Tulving, E., & Schacter, D. L. (1990, January 19). Priming and human memory systems. *Science, 247,* 301–306.

Vriezen, E., Moscovitch, M., & Bellos, S. (1995). Priming effects in semantic classification tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 933–946.

Whittlesea, B. W. A., & Jacoby, L. L. (1990). Interaction of prime repetition with visual degradation: Is priming a retrieval phenomenon? *Journal of Memory and Language, 29,* 546–565.

Woltz, D. J. (1990). Repetition of semantic comparisons: Temporary and persistent priming effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 392–403.

Woltz, D. J. (1996). Perceptual and conceptual priming in a semantic reprocessing task. *Memory & Cognition, 24,* 429–440.

Appendix A

Simulation Details

Pattern Presentation

For each pattern presentation, during both learning and testing–priming sessions, the states of the units were updated repeatedly for a number of time steps until the network reached a stable attractor. At each time step, t, each unit's state, $y_i(t)$, was updated according to the following equation: $y_i(t) = y_i(t-1) + \alpha[\sigma(x_i(t)/T) - y_i(t-1)]$, where α is fixed state adaptation rate parameter, $\sigma(x_i)$ is the activation function—a nonlinear function of the weighted summed input $x_i = \Sigma_j w_{ij} y_j(t)$—and T is the "temperature" parameter of the DBM. In our simulations, α was 0.4, and for the activation function we used the hyperbolic tangent nonlinearity,

$$\sigma(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}},$$

an S-shaped function producing real-valued continuous activations ranging from -1 to 1. In the learning and energy equations (described below), we shifted and scaled the activations to range from 0 to 1 so that they could be interpreted as probabilities. This function is commonly used in connectionist models with nonlinear units. Another common choice of nonlinearity is the sigmoid function, $\sigma(x) = 1/(1 + e^{-x})$, which is nearly identical in form. The temperature parameter scales the total input to units and thereby determines the slope of the activation function. Its purpose is explained in the next paragraph.

The Boltzmann machine is similar to a Hopfield (1982, 1984) network in that both use symmetric connections and Hebbian learning rules to store patterns as attractors. Two major problems with the Hopfield network are its poor storage capacity and the fact that it can settle into "blend states"—attractors that do not correspond to any stored patterns. The Boltzmann machine overcomes these two problems respectively by its learning rule, described in the next subsection, and by using a procedure called *simulated annealing* (Kirkpatrick, Gelatt, & Vecchi, 1983) to update the states until a low-energy attractor is reached. Simulated annealing works as follows: Each time a pattern is presented to the network, the unclamped units are initialized to intermediary states of 0.0, and the temperature parameter is set to a high initial value; at this high temperature, the slope of the activation function is nearly flat so units' states are relatively insensitive to their total inputs. Units' states are then gradually updated for many iterations. At each iteration, T is decayed by some factor, until it has reached a minimum value. This annealing procedure helps prevent the network from becoming trapped in poor local minima. In our simulations, the initial temperature was 30, the decay factor was 0.9, and the final temperature was 1.0. State updates continued until no unit had changed its state by more than 0.00001 from one iteration to the next. These parameter values were determined

empirically to produce reasonably fast and continuous convergence to an energy minimum.

Soft-clamping of the orthographic units' states was achieved by adding an extra connection with a large fixed weight to each orthographic unit and passing the external input in along this connection. The weights on these extra connections were fixed at 2.0, so that each orthographic unit's state would be influenced predominantly by the external input but also to some degree by the activations of units in the other layers in the network. We did not experiment with this choice of fixed weight. However, during learning the network was free to make the other weights arbitrarily large, so the choice of the external weight value became less relevant as learning proceeded.

Learning

In a standard DBM, the weight update for each connection is computed according to a contrastive Hebbian learning rule (Peterson & Anderson, 1987), which involves two phases for each training pattern. In the positive phase, the states of both the input and output units are clamped to their correct values, the input vector and desired output vector or training signal, respectively. The hidden unit states are then repeatedly updated until they settle to stable values. Positive Hebbian learning then occurs: $w_{ij}(t) = w_{ij}(t-1) + \epsilon y_i(t)y_j(t)$, where ϵ is the learning rate. Thus, weights on connections between strongly coactive units are strengthened. In the negative phase only the input units' states are clamped; the rest of the network settles into an attractor state that may or may not be the correct response, and this response is "unlearned": $w_{ij}(t) = w_{ij}(t-1) - \epsilon y_i(t)y_j(t)$. Thus, weights on connections between strongly coactive units are weakened. Eventually, learning in the two phases converges to a net effect of zero once the network has learned to produce the correct output states in response to the corresponding clamped input states in the negative phase. For efficiency of learning, the network was first trained for 2,000 sweeps through the training set of 40 patterns, exactly in the manner described by Plaut and Shallice (1993a), using hard-clamping of inputs. The learning rate for each unit during this initial training phase was $.01/fanin_i$, as in the Plaut and Shallice simulations (David Plaut, personal communication), where $fanin_i$ is the number of incoming connections to the ith unit. Thus, each unit's learning rate was scaled by its fanin. For each training pattern, the network was trained to map from orthography to the correct semantics and phonology, from semantics to the correct orthography and phonology, and from phonology to the correct semantics and orthography. Thus, each learning iteration involved one positive phase, hard-clamping the O, S, and P layers, and three negative phases, leaving the O and S layers, the P and S layers, or the O and P layers unclamped. The network was then trained for an additional 1,500 sweeps through the training set by using soft-clamping, with a fixed learning rate for all units of 0.00001, and this time it was only trained to map from orthography to semantics and phonology. During this latter training stage, zero-mean noise with a standard deviation of 0.05 was added to the orthographic inputs on each pattern presentation.

Free Energy

Hinton (1989) showed that the DBM learning procedure performs minimization by steepest descent in the Boltzmann free energy, F, which is defined in terms of the energy E and entropy H of the system as follows:

$$F = E - TH$$

$$= -\frac{1}{2}\sum_{ij} y_i y_j W_{ij} - T\sum_i [y_i \log y_i + (1 - y_i)\log(1 - y_i)].$$

It is this quantity that we plotted in Figure 2, where i and j indexed over all pairs of connected units in the entire network.

Appendix B

Hinton and Shallice's 40 Words

Indoor objects	Animals	Body parts	Foods	Outdoor objects
Bed	Bug	Back	Bun	Bog
Can	Cat	Bone	Ham	Dew
Cot	Cow	Gut	Hock	Dune
Cup	Dog	Hip	Lime	Log
Gem	Hawk	Leg	Nut	Mud
Mat	Pig	Lip	Pop	Park
Mug	Ram	Pore	Pork	Rock
Pan	Rat	Rib	Rum	Tor

Note. From "Lesioning an Attractor Network: Investigations of Acquired Dyslexia," by G. E. Hinton and T. Shallice, 1991, *Psychological Review, 98,* p. 79. Copyright 1991 by the American Psychological Association. Reprinted with permission of the author.

(*Appendixes continue*)

Appendix C

Hinton and Shallice's Semantic Features

Semantic features		
max-size-less-foot	max-size-foot-to-two-yards	max-size-greater-two-yards
main-shape-2D	main-shape-3D	
cross-section-rectangular	cross-section-circular	
has-legs		
white	brown	green
color-other-strong	varied-colors	transparent
dark	hard	soft
sweet	tastes-strong	
moves		
indoors	in-kitchen	in-bedroom
in-living-room	on-ground	on-surface
otherwise-suported		
in-country	found-woods	found-near-sea
found-near-streams	found-mountains	found-on-farms
part-of-limb	surface-of-body	interior-of-body
above-waist		
mammal	wild	fierce
does-fly	does-swim	does-run
living	carnivore	
made-of-metal	made-of-wood	made-of-liquid
made-of-other-nonliving	got-from-plants	got-from-animals
pleasant	container	
man-made	container	
for-cooking	for-eating-drinking	for-other
used-alone	for-breakfast	for-lunch-dinner
for-snack	for-drink	
particularly-assoc-child	particularly-assoc-adult	used-for-recreation
human	component	

Note. From "Lesioning an Attractor Network: Investigations of Acquired Dyslexia," by G. E. Hinton and T. Shallice, 1991, *Psychological Review, 98,* p. 94. Copyright 1991 by the American Psychological Association. Adapted with permission of the author.

Appendix D

Stimuli for Experiments 1 and 2

Sublist A				Sublist B			
Prime	Target	Prime	Target	Prime	Target	Prime	Target
whale	shark	cotton	silk	minute	hour	auger	drill
dolphin		rayon		second		borer	
fish		linen		week		chisel	
porpoise		satin		month		reamer	
salmon		velvet		day		pliers	
foot	mile	coat	shirt	waterway	river	ruby	jade
inch		pullover		stream		emerald	
yard		sweater		channel		sapphinre	
meter		dress		brook		topaz	
kilometer		blouse		creek		amethyst	
cherry	grape	stool	chair	knife	spoon	tulip	rose
plum		seat		fork		daisy	
lime		bench		ladle		violet	
raisin		recliner		silverware		orchid	
melon		throne		cutlery		pansy	
mansion	home	shuttle	train	cow	sheep	portal	door
house		railway		lamb		entranceway	
abode		freighter		goat		gateway	
dwelling		subway		mutton		window	
lodging		streetcar		mule		closet	
robin	dove	louse	roach	aluminum	steel	whiskey	beer
swan		flea		copper		pilsner	
pigeon		hornet		iron		stout	
eagle		parasite		zinc		lager	
canary		vermin		chrome		ale	
lentil	bean	manuscript	book	dollar	dime	knife	spear
broccoli		novel		nickel		skewer	
pea		journal		quarter		dagger	
legume		paperback		penny		sword	
spinach		document		coin		lance	
flurry	snow	salt	spice	zither	harp	evergreen	pine
blizzard		pepper		lute		birch	
sleet		seasoning		guitar		spruce	
slush		flavoring		cello		cedar	
frost		savory		piano		balsam	
pontoon	raft			paw	hand		
surfboard				foot			
buoy				mitt			
float				palm			
ferry				finger			

(*Appendixes continue*)

Appendix E

Stimuli for Experiment 3

	Stimulus pairs					
Practice stimuli	Prime	Target	Prime	Target	Filler words	
board	fate	destiny	snare	trap	stuff	fun
butter	infant	baby	demon	devil	blood	mildew
lettuce	foam	froth	tack	pin	ant	skunk
chasm	iceberg	glacier	plate	dish	bear	carrot
cabin	wasp	hornet	bison	buffalo	fight	moss
crab	mistake	error	rug	carpet	aardvark	goldfish
bacteria	sorcery	magic	rifle	gun	muffin	young
cord	filth	dirt	bucket	pail	mushroom	cactus
priest	pig	swine	book	novel	shark	mule
frost	cushion	pillow	mist	fog	paste	gopher
hall	sound	noise	prison	jail	rose	cucumber
jury	odor	scent	home	house	robin	parrot
gem	marsh	swamp	monkey	ape	stone	folder
officer	frog	toad	ladle	spoon	apple	pill
oven	kleenex	tissue	desk	table	rope	frame
person	agony	pain	wolf	dog	germs	trout
lizard	fire	flame	oyster	clam	fly	lobster
shadow	bush	shrub	boat	ship	grape	iguana
salute	robber	thief	hawk	eagle	weed	bottle
square	shore	beach	mouse	rat	under	shelf
	melody	song	snake	serpent	coral	clock
	octopus	squid	student	pupil	soap	kite
	shrine	altar	grass	lawn	chief	bone
	pledge	promise	sheep	lamb	doctor	daisy

Received June 7, 1995
Revision received March 3, 1997
Accepted March 3, 1997 ■