Journal of Memory and Language **46,** 24–38 (2002)
doi:10.1006/jmla.2001.2802, available online at http://www.academicpress.com on IDEAL®

Can Distributed Orthographic Knowledge Support Word-Specific Long-Term Priming? Apparently So

Jeffrey S. Bowers, Markus F. Damian, and Jelena Havelka

Department of Experimental Psychology, University of Bristol, Bristol, England

A property of distributed representations is that related information is coded as overlapping patterns of activation over the same set of units and learning associated with one item extends to related items. Accordingly, the null (or near null) long-term priming observed between form-related words seems to pose a challenge to connectionist theories of reading that include distributed codes. In the present report, priming was assessed in a behavioral study and a computer simulation using Seidenberg and McClelland's (1989) distributed model of word identification. Contrary to our expectation, both the behavioral and simulation studies obtained robust repetition and little form priming. Furthermore, analysis of the model's performance revealed that the lack of form priming was the product of collapsing facilitatory effects between rhymes (*mint–hint*) and inhibitory effects between nonrhymes (*pint–hint*). A second behavioral experiment confirmed this prediction. A number of additional long-term priming results were also successfully modeled. © 2001 Elsevier Science

Key Words: distributed representations; localist representations; long-term priming; form priming; word identification.

One key difference among theories of single-word reading concerns the structure of orthographic knowledge. According to early models by Morton (1979), Forster (1976), and McClelland and Rumelhart (1981) and more recent models by Coltheart, Curtis, Atkins, and Haller (1993), Grainger and Jacobs (1996), Norris (1994) and others, words are represented in a localist fashion. That is, words are coded separately from one another with a discrete lexical-orthographic representation for each word (or root morpheme) in a person's vocabulary. By contrast, according to most connectionist models, words are represented as a distributed pattern of activation over a collection of units, with the same set of units contributing to the representations of many words (McClelland & Rumelhart, 1985; Seidenberg & McClelland, 1989; for exception see Grossberg & Stone 1986; for general discussion of localist connectionism see Grainger & Jacobs, 1998). On this view, there are no discrete lexical-orthographic codes corresponding to specific words.

At present, models with localist or distributed orthographic (and phonological) representations can account for dozens of word and nonword naming results in both skilled readers and persons with various forms of acquired dyslexia (e.g., Plaut, McClelland, Seidenberg, & Patterson, 1996). Although localist models might appear better suited to accommodate various lexical decision results, some recent distributed models show promise in accommodating these findings as well (e.g., Plaut, 1997). Thus, there are no strong grounds for favoring one approach over another when considering these results. Unlike most localist accounts, connectionist theories incorporate learning mechanisms and thus have the potential to explain how orthographic and other forms of knowledge are acquired in the first place. By exploiting these learning principles, connectionist models can accommodate various findings concerning the development of reading skills in both normal and dyslexic children (e.g., Harm & Seidenberg, 1999). These latter issues fall outside the domain of most theories with localist representations (but see Davis, 1999).

In the present investigation we explore a phenomenon that has attracted relatively little atten-

We thank Ken Seergobin who provided us with the trained S&M(89) model as well Eric Wright for assisting us with his programming skills. We also thank Mike Masson, Art Jacobs, and an anonymous reviewer for their helpful comments.

Address correspondence and reprint requests to Jeffrey Bowers, Department of Experimental Psychology, University of Bristol, 8 Woodland Road, Bristol BS8 1TN, United Kingdom. E-mail: j.bowers@bris.ac.uk.

0749-596X/01 $35.00

tion from either camp; namely *long-term priming* for words (but see Becker, Moscovitch, Behrmann, & Joordens, 1997; McClelland & Rumelhart, 1985; Morton, 1979; Stark & McClelland, 2000). Long-term priming refers to a facilitation in the processing of a stimulus as a consequence of encoding the same or a related stimulus in an earlier episode. For example, priming in the lexical decision task is observed when participants are faster and more accurate in categorizing letter strings as words (as opposed to nonwords, such as *blap*) when they were studied earlier. This priming is called long-term because it lasts minutes, hours, and sometimes longer, distinguishing it from various sorts of short-term priming, such as masked or semantic priming that typically last only a few seconds (for a connectionist accounts of short-term priming, see Masson, 1995; Plaut & Booth, 2000; but see Becker et al., 1997, for evidence that semantic priming can persist beyond a few seconds under some conditions). Although it is sometimes argued that long-term priming is mediated by episodic memory representations that are separate from lexical-orthographic representations (e.g., Forster & Davis, 1984), there is now strong evidence that priming for written words is largely a by-product of learning within the orthographic and phonological systems (e.g., Bowers, 1999, 2000a; Bowers & Michita, 1998). On this later view, long-term priming should provide constraints to theories of word recognition—particularly those concerned with issue of learning.

Accordingly, we attempted to simulate long-term priming using a connectionist model of word identification that learns distributed word representations via the back-propagation learning algorithm—representational and learning assumptions shared by most connectionist models of reading (e.g., Seidenberg & McClelland, 1989; Plaut et al., 1996; Harm & Seidenberg, 1999). Long-term priming has not previously been simulated in these models, and it is not at all obvious that they will succeed. Learning with back-propagation is subject to a phenomenon called “catastrophic interference” in which new information erases old information (Grossberg, 1987; McClosky & Cohen, 1989; Ratcliff, 1990). In order to avoid this interference so that large vocabularies can be acquired, it has been necessary to reduce the learning rates and to introduce an interleaved study regimen in which all the words in the vocabulary are acquired in parallel. This solution works well for many purposes, but it raises the question of whether this slow learning can support long-term priming that occurs following a single-study trial. Ratcliff and McKoon (1997) argued that the Seidenberg and McClelland (1989) model was incapable of supporting long-term priming because of its slow learning rates. Indeed, connectionist models of memory that have assessed long-term priming tend to use different learning rates during the training of the network and the critical priming trials (e.g., Becker et al., 1997; McClelland & Rumelhart, 1985). Even if the learning rates used for training are found to be sufficient to support priming for a number of learning trials, the interference effects due to learning unrelated words processed between study and test may be incompatible with the longevity of priming. Interestingly, all past connectionist models that have assessed long-term priming include small vocabularies, such as 48 words in the recent Stark and McClelland (2000) study.

In addition to these potential problems, a number of related long-term priming results appear to pose a challenge for these models. For example, little or no priming is obtained between form-related words, such as *card/car* (e.g., Napps & Fowler, 1987; Ratcliff & McKoon, 1997; Rueckl & Mathew, 1999). This result appears problematic for models with distributed representations because the orthographic representations of form-related words (e.g., *card* and *car*) overlap. Thus, any learning that occurs for *card* should impact *car* as well as *card*. Indeed, these models depend on learning between form-related items in order to account for a number of key results in the reading literature, including the interaction between frequency and consistency in naming latencies (Seidenberg & McClelland, 1989). Thus, it might be expected that any repetition priming observed in these models would be associated with form priming effects. Form-specific priming results seem more consistent with theories that include separate localist representations (e.g., Morton, 1979).

Our main goal was to determine whether the pattern of robust repetition and near null form priming is consistent with the general principles of distributed word representations and learning via back-propagation. To this end, we assessed priming using the Seidenberg and McClelland (1989) model that was trained with the same set of words and learning parameters originally employed; henceforth, the S&M(89) model. The key feature of this model for our purposes is that information is coded in a distributed fashion at all levels, from the orthographic input units to the phonological output codes, and it learns by back-propagation. Note that we are not trying to evaluate the adequacy of the S&M(89) model as a model of reading or priming. A number of important limitations have been identified, including its poor performance in reading nonwords and its inability to simulate various acquired reading disorders (Besner, Twilley, McCann, & Seergobin 1990; Coltheart et al., 1993). Instead, we are asking the more general question of whether the back-propagation and distributed coding schemes used in a variety of models are compatible with word-specific priming effects documented in the literature. For this purpose, the S&M(89) model actually provides a stronger test than the next generation connectionist model introduced by Plaut et al. (1996) and Harm and Seidenberg (1999), which includes sublexical localist coding schemes in the orthographic input and phonological output layers.

In order to directly compare the simulated priming results to empirical data, we first carried out a behavioral study of repetition and form priming (Experiment 1), which allowed us to present the same set of words to the model. As we demonstrate, the model does an excellent job in simulating the behavioral results. Second, in analyzing the performance of the model, we discovered that the model makes a novel prediction regarding form priming; namely facilitatory form-priming should be obtained between words that rhyme (*mint–hint*) and inhibitory form-priming should be obtained between nonrhymes (*pint–hint*). We tested this prediction in a behavioral experiment (Experiment 2), which verified the prediction. Finally, the model was tested on a number of additional variables, including its ability to accommodate the strong interaction between priming and frequency, with reduced priming for high-frequency words (e.g., Bowers, 2000b), and the longevity of priming, with priming lasting days or months under some conditions (e.g., Sloman, Hayman, Ohta, Law, & Tulving, 1988). Again, the model performed well. Based on these results, we conclude that connectionist models that learn distributed representations via back-propagation can accommodate various aspects of long-term priming. More generally, we conclude that the large literature in long-term priming should be used to constrain models of word identification.

EXPERIMENT 1

A variety of studies have shown little or no long-term priming between form-related words. For example, Ratcliff and McKoon (1997) assessed form priming in four perceptual identification experiments and, averaging across experiments, obtained a robust repetition effect (15% improvement above baseline) and no form priming (1% improvement). Similarly, averaging across two experiments, Rueckl and colleagues (Rueckl et al., 1997, Experiment 2; Rueckl & Mathew, 1999; Experiment 4) obtained robust repetition effects (17% improvement) and no form priming (2% improvement) in the stem and fragment completion task (for similar results, see Hanson & Wilkenfeld, 1985; Murrel & Morton, 1974). Indeed, Napps and Fowler (1987) obtained a robust repetition effect (45-ms reduction in RT latencies and 5% improvement in accuracy) and a trend for inhibitory form priming (11-ms increase in latency and 1% reduction accuracy) in the lexical decision task. Significant (albeit small) form priming has been obtained following multiple study trials (e.g., Rueckl, 1990), and averaging across all studies, there is a general trend for form priming following a single study trial. Thus, there may well be a small effect.

Because none of the above studies restricted their stimuli to single-syllable words, we carried out a form priming study with a set of single-syllable words included in the vocabulary of the original S&M(89) network. Because we wanted to focus on the orthographic contributions to long-term priming, we included procedures in the experiment intended to maximize the orthographic and minimize the phonological contributions to priming. In particular, we assessed priming using the lexical decision task and included pseudohomophones (e.g., *brane*) as the nonword foils. Under conditions in which all the nonwords sound like real words, there is some evidence that phonological codes contribute less to the process of making lexical decisions (e.g., Andrews, 1982; Davelaar, Coltheart, Besner, & Jonasson, 1978; Stone & VanOrden, 1993; but see Grainger & Jacobs, 1996; Pexman et al., 1996). However, it is important to emphasize that orthographic codes largely mediate visual word priming even when pseudowords (e.g., *blap*) are included as the nonword distracters (e.g., Bowers, 2000b; Bowers & Michita, 1998).

Method

Participants. Forty-eight students from the University of Bristol participated in return for course credit or payment.

Design and materials. Fifty-six single-syllable word pairs that were orthographic neighbors (e.g., *crab–crib*) and that were included in the original training set of the S&M(89) model were selected. One item from each pair was randomly selected as the target (mean frequency = 14, range = 0–320 occurrences per million according to the CELEX Lexical Database), the other as the prime (mean frequency = 24, range = 1–905; Baayen, Piepenbrock, & van Rijn, 1993). A further set of 56 single-syllable pseudohomophones served as the nonword foils (see Appendix A for list of items). At study, participants were presented with 14 visually presented targets (repeated condition), 14 auditorily presented targets (cross-modal condition), and 14 visually presented primes (form priming condition), while 14 words were not presented (baseline condition). At test, all 56 targets were presented visually. Four test forms were constructed in order that all words were presented in all conditions.

Procedure. The experiment was conducted under conditions of incidental encoding: Participants were told that they were participating in an experiment concerned with word perception, and they were not informed that items presented at study were later repeated at test. During the study phase, spoken and lowercase written words were presented every 3 s in a random order. To insure that participants were paying attention, they were required to press the left shift key on the keyboard for words with a negative connotation and the right shift key for words with a positive connotation. These responses were not recorded. Immediately following the study phase, participants performed the lexical decision task. The experiment included a set of 20 practice items (10 words and 10 pseudohomophones) that were different from the critical 56 test words and the 56 pseudohomophones that followed. On each trial, a fixation point (+) was displayed for 500 ms followed by the target displayed in lowercase letters for 500 ms. Participants were instructed to press the right shift key of the computer keyboard as quickly as possible if the item was a word and the left shift key for a pseudoword. Participants were informed that all the nonwords sounded like words. Items were presented in a different random order to each participant and were presented on a Multisync monitor controlled by a Pentium PC using the DMASTR display software developed by Kenneth Forster and Jonathan Forster at the University of Arizona. Standard IBM text font was used, and participants viewed the screen from approximately 50 cm.

Results and Discussion

The response latencies and error rates in the various conditions are presented in Table 1. For the response times, repetition priming (21 ms) was significant, $t1(47) = 2.07$, $p < .05$; $t2(55) = 2.97$, $p < .01$, whereas no significant effects were obtained in the cross-modal (6 ms) nor form (1 ms) priming conditions, both $t1(47) < 1$; $t2(55) < 1$. For errors, repetition priming (5.5%) was highly significant, $t1(47) = 2.75$, $p < .01$; $t2(55) = 3.34$, $p < .01$, whereas the cross-modal (2.5%) and form (1.3%) priming were not, both $t1(47) < 1.14$, $p > .25$ and $t2(55) < 1.38$, $p > .15$. Thus, consistent with past research, repetition priming was obtained in the

TABLE 1

Response Latencies (in Milliseconds) and Percentages of Error Rates in the Lexical Decision Task as a Function of the Priming Conditions in Experiment 1 and the Associated Estimated Latencies for the S&M(89) Model

Prime conditions	Experiment RTs	Experiment Errors	Simulation Estimated RTs
Repeat	631 (21)	7.1 (5.5)	626 (24)
Cross-modal	646 (6)	10.1 (2.5)	—
Form-prime	651 (1)	11.3 (1.3)	649 (1)
Baseline	652	12.6	650

Note. Priming scores are in parentheses.

context of little or no form priming. This word-specific priming can be attributed, in part, to orthographic representations, given that the cross-modal priming was not significant. However, the inclusion of the pseudohomophones as distracter foils was only partially successful in restricting priming to the orthographic system given that cross-modal priming approached one third the size of repetition priming.

Simulating repetition and form priming. As noted above, we used the S&M(89) model to assess form and repetition priming. The model includes an orthographic layer composed of 400 units, with each unit coding for 1000 different letter triplets such that the activation of a single orthographic unit is highly ambiguous. However, each letter triplet in a word (i.e., WOR or ORD in WORD) activates approximately 20 input units such that the pattern of activation across all activated units for a given word uniquely identifies the word. The orthographic codes connect with a set of 200 hidden units which feedback onto the orthographic units as well as connect with a set of 460 phonological output units. Similar to the orthographic layer, each phonological unit in the model represents many different combinations of phoneme triplets, and each phoneme in a word activates many different units, with the pattern of activation across the units uniquely specifying the phonology of the word. Orthographic processing in the model consists in transforming an input into a pattern of activation over the hidden units, which in turn feed back onto the orthographic codes, and naming consists in transforming patterns of activation in the orthographic layer into patterns of activation in the hidden layer, followed by activation in the phonological layer. See S&M(89) for more details of the model.

In order to assess priming in the S&M(89) model in an analogous fashion to the behavioral study, we carried out 48 separate simulations, with each simulation corresponding to a single participant. In the study phase of each simulation, the model was presented with the 14 words taken from the repeated and form-related conditions. Following each study trial, the connection weights were modified to the same degree as the learning trials on which it was originally trained. As in the behavioral study, words were rotated through the repeated and form-related conditions, but words that had been assigned to the auditory study condition in the behavioral study were not presented to the model (as the model was not designed to encode spoken words). Then at test, all 56 words were presented, and the orthographic error score associated with each word was computed. That is, the sum of the squared differences between the target activation value for each orthographic unit and its actual activation was computed. The error score reflect the extent to which feedback from the hidden unit was successful in reconstructing the input pattern. Again, see S&M(89) for details. Four study lists were constructed in order to achieve full counterbalancing, and items were presented in a different random order in each simulation. Priming was computed by comparing the orthographic error scores for words in the repeated and form-related conditions compared to the baseline condition, averaging across all simulations. Thus, unlike the behavioral study, nonwords were not presented to the model in the test phase.

In order to easily relate the error score results to the reaction time measures reported in the behavioral study, we transformed the error scores to RTs, using the formula described by S&M(89, p. 532). The authors estimated empirical latencies to be approximately 10 times the error score plus a constant of 500–600 ms (we used 550 ms). These estimates were not derived to fit long-term priming data, but were based on various naming and lexical decision studies the authors reported. Nevertheless, Table 1 shows the RT estimates for repetition priming were similar to the behavior results. This outcome suggests that the learning rates employed in the S&M(89) model are appropriate not only to support the learning of distributed representations but also to support long-term priming following a single study episode.[1]

The pattern of priming in the S&M(89) simulation mirrored the pattern of priming in our participants, with the estimated priming scores in the model showing dramatically reduced form (1 ms) compared to repetition (24 ms) priming (see Table 1). Consistent with this analysis, a series of planned contrasts revealed highly significant repetition priming, $t1(47) = 14.84$, $p < .001$, $t2(55) = 11.74$, $p < .001$, and a form priming effect that only approached significance, $t1(47) = 1.56$, $p = .126$, $t2(55) = 2.14$, $p = .037$. The reason a 1-ms priming effect approached significance is that the S&M(89) model is deterministic and the only variability across simulations was due to the fact that words were presented in different random orders in each counterbalanced file. In any case, the S&M(89) model with its distributed representations can account for the null (or close to null) form priming effects coupled with robust repetition priming.

Although the model can accommodate this pattern of repetition and form priming, the question remains as to why it performs the way it does. Part of the answer was in fact suggested by S&M(89), who assessed *short-term* priming on a single target word (*tint*) that was immediately preceded by a form-related word that rhymed (*mint*) or that did not (*pint*). Measuring the outputs from the phonological units, they observed small positive priming between the rhyming pair and a small negative priming between the nonrhyming pair. A combination of facilitation and inhibition may help explain the small form priming effects we obtained, given that rhyme and nonrhyme pairs were included in our experiment.

In order to check for this possibility, we compared form priming for the rhyme and nonrhyme pairs, but in this case considered the outputs of the orthographic units. Forty-three of our items rhymed, and they showed an estimated orthographic facilitation of 2.6 ms. Thirteen of our pairs did not rhyme and they showed similar sized inhibitory effect of 2.3 ms, both highly significant. Clearly, our long-term orthographic priming effects mirrored the short-term phonological priming effects reported by S&M (89).[2]

Still, it is important to emphasize that the facilitation was only an estimated 2.6 ms for the rhyme items compared to the repetition effect of 25 ms for the same items (the non-rhyming items showed a repetition effect of 22 ms). Thus, these inhibitory form priming effects for the nonrhyme items appear to play only a small

---

[1] One limitation of the S&M(89) model is that the only measure of orthographic processing is the orthographic error score. Accordingly, it is not possible to make independent estimates of RT latencies and error rates. Given that the reduction in the orthographic error scores was used to make estimates in improved processing in the RT measures and given that the behavioral studies obtained priming in the response latencies and error scores, it is reasonable to conclude that the overall magnitude of priming in the model was reduced compared to the behavioral results. But given that we did not change any of the parameters of the model, this is not surprising. The important point is that the pattern of priming in the model and in humans matched, and this is not compromised by a failure to separately measure error scores.

[2] The effect of rhyme status on orthographic error scores might appear to be a counterintuitive finding, but it makes sense when one realizes that orthographic representations within the model were reorganized during the process of learning the orthographic–phonological correspondences. That is, during the learning of these correspondences, the back-propagation algorithm changed connection weights between both the phonological output units and the hidden units (which reside outside the orthographic system in this model) and between the orthographic input units and the hidden units (which constitute the orthographic system within the model). Thus, in this model, orthographic knowledge was not only organized in a bottom-up fashion according to the statistical regularities of the visual inputs, but was also organized in a top-down fashion from a phonological “teacher.” It was the influence from the phonological teacher that caused rhyme status to affect orthographic error scores.

Note, there are good reasons to argue that orthographic knowledge is affected by a phonological teacher in skilled readers as well. For instance, there is now strong evidence that orthographic knowledge is coded in an abstract format, with visually dissimilar exemplars of letters and words mapping onto abstract letter and word codes (e.g., a/A and role in explaining the small overall form priming effects. Apparently, the hidden units employed in the model learn to encode different words relatively independently of one another, consistent with the general analysis of Hanson and Burr (1990) concerning the role of hidden units in connectionist networks.

EXPERIMENT 2

Despite the small facilitatory and inhibitory priming effects obtained for the rhyme and nonrhyme items, the simulation raises the interesting possibility that a similar effect would be observed in a behavioral study. As an initial test of this prediction, we checked whether this pattern was obtained in the behavioral data of Experiment 1. The differences were striking, with a 54-ms inhibitory pattern for the 15 items that did not rhyme and an 17-ms advantage for the 41 items that did. We do not want to make too much of these findings given that the analysis was post hoc, so we carried out a more systematic test of this prediction. In Experiment 2, we selected word triplets such that one item (the target, e.g., *hint*) rhymed with a form-related prime word (*mint*) and did not rhyme with the other form-related prime (*pint*). We asked whether form-related priming in the lexical decision task was modulated by the rhyme status of the prime–target pairs.

Method

Participants. Fifty-two students from the University of Bristol participated in return for course credit or pay.

Design and materials. Twenty-eight triplets of form-related words were selected such that they all differed in the first letter position. Two items from each triplet rhymed, whereas the third item did not. One of the rhyming items was randomly selected as the target, such that the target could be primed with itself (repetition condition), with a form-related rhyme (rhyme condition), or a form-related nonrhyme (nonrhyme condition). The mean frequency of the target, rhyme, and nonrhyme items were 10, 15, and 15 occurrences per million (Baayen et al., 1993), with ranges of 0–53, 0–70, and 1–56, respectively. See Appendix B for the list of words. A further set of 28 single-syllable pronounceable nonword foils was selected. The inclusion of pseudowords as opposed to pseudohomophones makes the present study more similar to past form priming experiments. During the study phase, seven prime words were presented in each of the repetition, rhyme and nonrhyme conditions. At test, all 28 targets were presented. Four test forms were constructed in order to achieve complete counterbalancing.

Procedure. The experiment was conducted under conditions of incidental encoding. During the study phase, lowercase written words were presented every 2 s in a random order. In order to insure that participants were paying attention, they were required to say each item aloud. Immediately following the study phase, participants were given instructions to perform the lexical decision task. On each trial, a fixation point (+) was displayed for 500 ms followed by the target displayed in lowercase letters for 500 ms. Participants were instructed to press the right shift key of the computer keyboard as quickly as possible if the item was a word and the left shift key for a pseudoword. Items were presented in a different random order to each participant and were displayed on a Multisync monitor controlled by a Pentium PC using DMASTR.

---

read/READ map onto a abstract visual codes for *a* and *read*, Bowers, 1996; for a review, see Bowers, 2000a). More strikingly, abstract orthographic codes also develop between the visually unrelated Japanese scripts of Hiragana and Kanji (Bowers & Michita, 1998), and there is no way in which these latter mappings can be learned on the basis of the visual structure of the Japanese writing systems. Based on these findings, the first author has argued that some sort of top-down influence was required to learn these arbitrary mappings and concluded that phonology (and perhaps semantics) acts as a teacher (albeit in a different way than the teacher in the S&M(89) model). Other findings that the orthographic system maps together morphologically related words (e.g., Rapp, 1992), or represents words in terms of onsets and rhymes (e.g., Treiman, Mullennix, Bijeljac-Babic, & Richmond-Welty, 1995) may also reflect top-down influences from semantics and phonology on structuring orthographic knowledge. Consistent with this conclusion, various evidence suggests that there are bidirectional connections between orthography and phonology such that active phonological codes lead to the automatic activation of orthographic codes and vice versa (e.g., Stone, Vanhoy, & Van Orden, 1997; but see Peereman, Content, & Bonin, 1998).

Results and Discussion

The response latencies and error rates in the various conditions are presented in Table 2. Planned comparisons carried out on the RT latencies confirmed that repetition (49-ms) and rhyme (31-ms) priming were significant, with both $t1(48) > 2.57$ and $t2(24) > 2.24$, $p < .05$, whereas nonrhyme (-2 ms) priming did not approach significance, with $t1$ and $t2 < 1$. The analysis on the error scores showed that the positive repetition priming (4.6%) approached significance, $t1(48) = 2.86$, $p < .05$; $t2(24) = 1.79$, $p = .09$, as did the inhibitory nonrhyme priming (-3.5%), $t1(48) = 1.79$, $p = .08$; $t2(24) = 1.65$, $p = .11$. An inhibitory rhyme effect (-0.3%) did not approach significance, with $t1$ and $t2 < 1$. As the model predicted, form priming was facilitatory for items that rhymed and inhibitory for items that did not. If the rhyme and nonrhyme items were collapsed (as is common in many past priming experiments, including our Experiment 1), the typical null form priming effect would have been obtained, with a 15-ms facilitatory RT priming effect countered by a 1.8% inhibitory priming effect in errors.

What is surprising, however, is that the facilitatory and inhibitory priming effects for the rhyme and nonrhyme items was much larger in the behavioral compared to the simulation study. Whereas form priming for the rhyme and nonrhyme items was approximately 10% in the model, it was roughly 30% in the behavioral study (when considering both the RTs and error rates). One possible explanation for this discrepancy is that the form priming results in the model were based solely on the orthographic error scores, whereas our participants may have relied on both orthographic and phonological representations when making lexical decisions (the pseudohomophone distracter foils were not used in this task). Accordingly, we calculated priming effects for the rhyme and nonrhyme items based on the phonological error scores in our simulation. These priming scores were even smaller, with a 2.0 ms facilitatory and -0.1 ms inhibitory effect, respectively. If the priming effects calculated from the orthographic and phonological error scores are combined, the differences would be reduced, but even then priming in the model is more word-specific than the priming obtained in our behavioral study. Another possible explanation for the discrepancy is that the form priming effects we obtained in the behavioral study are overestimates and that the size of these effects will be reduced in future studies. This can only be determined with additional studies. Whatever the outcome of future studies, the simulation study predicted this unexpected pattern of priming. The results indicate that distributed word representations are not incompatible with word-specific priming data. If anything, priming with these representations can be too word-specific, exactly opposite to our initial intuitions.

TABLE 2

Response Latencies (in Milliseconds) and Percentages of Error Rates in the Lexical Decision Task as a Function of the Priming Conditions in Experiment 2

Prime conditions	RTs	Errors
Repeat	615 (49)	3.6 (4.7)
Rhyme	634 (30)	8.5 (−.2)
Nonrhyme	666 (−2)	11.6 (−3.3)
Baseline	664	8.3

Note. Priming scores are in parentheses.

Before discussing the more general implications of the findings, we describe a number of additional simulations in order to determine whether the general principles of distributed representations and back-propagation are consistent with the interaction of frequency and priming as well as the longevity of priming.

Simulating frequency effects in the S&M(89) model. As noted in the introduction, priming is smaller for high- than low-frequency words (e.g., Bowers, 2000b; Forster & Davis, 1984). Accordingly, we contrasted repetition priming for a set of 48 high- (mean frequency approximately 300 occurrences per million) and low- (mean frequency 3 occurrences per million) frequency words in the S&M(89) model (Baayen et al., 1993). Twenty-four of the high- and low-frequency items were presented at study, and all items were presented at test, with the nonstudied items constituting the baseline condition in order to determine priming. We carried out 48 separate simulations, with each simulation corresponding to a single participant; studied and nonstudied items were rotated across simulations. The estimated response latencies and priming scores of the model are presented in Table 3, along with the results averaged across three lexical decision studies from Bowers (2000b) that included a set of 48 words with similar frequencies. (Note: Because many of the words included in the behavioral experiments were two syllables in length, the simulation study did not include the same set of words.) Estimated response latencies were calculated here as 10 times the orthographic error score plus a constant of 480 ms. The different constant used in the two experiments is only for the purpose of matching the behavioral and simulation baseline RTs and does not contribute to the size of the priming estimates. As can be seen in the table, priming in the model and in the behavioral study were again similar, demonstrating that the interaction between priming and frequency can readily be accommodated within the model despite its rejection of lexical representations.

The durability of long-term repetition priming in the S&M(89) model. One striking feature of long-term priming is its persistence, with priming lasting a few hours (Squire, Shimamura, & Graf, 1987) or many months (Sloman et al., 1988), depending on the task, the subject population, and other factors. This raises the question of how persistent priming effects are in the S&M(89) model. To assess the durability of priming, we reran the first simulation experiment described above, except that the model was exposed to unrelated words following the study phase, with the same connection weight changes applied to these as to all other learning trials. The number of unrelated items presented to the model was varied across experiments, and items were randomly sampled from the entire vocabulary set. The repetition priming scores following 50, 100, 500, 1000, 2000, 4000, and 8000 intervening trials were 26.7, 23.8, 18.3, 13.7, 9.5, 7.2, and 3.5 ms, respectively. These results show that the priming effects are quite persistent, with some priming surviving 8000 learning trials (more than three times larger than the vocabulary of the model).

It would be interesting to translate these results into an estimate of the longevity of priming in the model, but this is difficult for a number of reasons. One problem is that the model and the typical participant in an experiment have very different-sized vocabularies, and thus it is not clear what 8000 learning trials should correspond to in the participant. Should we estimate the typical time it takes before a person reads 8000 words or the time it takes to read 3 × 30,000 words, a rough estimate of the number of words in a typical person's vocabulary (Levelt, 1989)? Another problem is that it is not clear whether words read by a person outside the laboratory are perceptually encoded to the same extent as words encoded at study list given that (a) most words outside the lab are read in the context of a meaningful sentence (such that many words can be inferred from minimal perceptual encoding) and (b) eye fixations to a given word in a sentence are approximately 200–250 ms, much less than the time spent encoding words in a study list (see Bowers, 2000a, for a more detailed discussion). Indeed, Subramaniam, Biederman, and Madigan (2000) recently reported a failure to obtain any priming for pictures pre-

TABLE 3

Response Latencies (in Milliseconds) and Error Rates (in %) in the Lexical Decision Task as a Function of Word Frequency, and the Associated Estimated Latencies for the S&M(89) Model.

		High frequency		Low frequency	
Experiment	Condition	RTs	Errors	RTs	Errors
Bowers (2000b)	Repeat	518 (8)	2.8 (.5)	553 (33)	6.3 (6.4)
	Baseline	526	3.2	586	12.7
Simulation	Repeat	514 (6)	—	565 (25)	—
	Baseline	520	—	590	—

Note. Priming Scores are in parentheses.

sented up to 31 times in an RSVP sequence when items were displayed for between 72 and 126 ms per picture. At the same time, the pictures could be identified at these durations. Based on these findings, the authors argued that priming requires participants to attend to an item for a period of time after the item has been identified. They cited the findings of Tovee and Rolls (1995), who found cells in the inferior temporal that fire in response to specific stimuli in the first 50 ms of stimulus presentation and continue their activity for an additional 350 ms. According to Subramaniam et al. (2000), the additional activity may be required for memory encoding, and this activity is disrupted by attention to the next image during RSVP presentation or, possibly, in the case of reading text, by the next fixated word.

Based on these considerations, it might be argued that the learning rate for the nonstudied items should have been reduced, as these items are intended to correspond to the words encountered outside the experimental setting, most of which would be read in text. Reducing the learning for these items would increase the persistence of long-term priming in our simulation. In any case, the present results show that the S&M(89) model can accommodate relatively long-lasting priming effects.

GENERAL DISCUSSION

Our key finding is that the S&M(89) model that learns distributed word representations via back-propagation supports long-term word-specific priming. These findings contradict Ratcliff and McKoon (1997), who claimed that the slow learning associated with back-propagation precludes robust long-term repetition priming following a single episode. The results also appear to contradict our original assumption that repetition and form priming should co-occur in any model that included distributed coding schemes. However, it turned out that the absence of form priming in the model was the product of collapsing small facilitatory form priming effects for items that rhyme (e.g., *hint–mint*) with small inhibitory form priming effects for items that do not rhyme (e.g., *pint–mint*). Accordingly, repetition and form priming do occur in the model, as we originally assumed. The same pattern of facilitatory and inhibitory form-priming was obtained in a behavioral study for rhyme and nonrhyme items, suggesting that past reports of null form-priming in the literature may be a consequence of collapsing these effects, at least in the lexical decision task.

In addition, we demonstrated that the interaction between frequency and priming in behavioral studies can be explained as a natural byproduct of learning with back-propagation and that these simulated priming effects can persist for thousands of trials. The S&M(89) model performed well despite the fact that it was not designed to accommodate long-term priming, we did not vary any of the parameters from the original model, and the first author has a theoretical commitment to localist coding schemes in visual word recognition based on other findings and considerations (Bowers & Michita, 1998; Bowers, 2000a, Bowers, submitted). We doubt there are many examples of models performing so well under similar conditions.

Although the S&M(89) model was remarkably successful at simulating various long-term priming phenomena, we are not attempting to support this particular model of word identification. As noted above, the model is unable to accommodate a number of phenomena, including nonword naming and acquired dyslexia. More recent versions of the model have been published to deal with many of these weaknesses (Harm & Seidenberg, 1999; Plaut et al., 1996; but see Besner, 1998; Bowers, submitted; Page, 2000). What is critical for our purposes is that later versions of the S&M(89) model also represent word knowledge in a distributed fashion and learn via back-propagation. Our goal was simply to show that these general representational and learning principles are consistent with long-term priming phenomena, something that was not at all clear before the present simulations.

Given recent evidence that long-term priming reflects a form of orthographic learning that improves processing of repeated words (Bowers, 1999), it will be important to test whether other models of word identification that adopt different representational and learning assumptions can also accommodate long-term priming.

A number of models with localist codes have been designed to accommodate various short-term priming effects (e.g., Forster & Davis, 1984; Grainger & Ferrand, 1996), but they have largely ignored long-term priming and, for the most part, do not include learning mechanisms that could support these effects. One recent attempt to account for long-term priming in a localist model of word identification was made by Ratcliff and McKoon (1997), who developed their *countermodel* of word identification around various long-term priming data. The model did not include any learning mechanisms, and as a consequence, it explained all priming as a function of bias that does not improve word processing, as did the Morton (1979) model. Although learning mechanisms for low-frequency words were included in a more recent version (McKoon & Ratcliff, in press) in order to account for evidence that priming facilitates processing of repeated low-frequency words (Bowers, 1999; Wagenmakers, Zeelenberg, & Raaijmakers, 2000), the model does not include general learning algorithms that can support the acquisition of orthographic knowledge. As a consequence, like most localist models, it needs to be handwired and cannot account for the fact that priming extends to nonwords (e.g., Bowers, 1996; Stark & McClelland, 2000). Both versions of the counter model are also unlikely to accommodate the present set of results given that it has no phonological representations that could modulate form priming for the rhyme and nonrhyme items. More generally, both versions of the countermodel have not been tested on their ability to accommodate a wide range of single-word reading phenomena and thus cannot be considered competitors to existing models of visual word recognition. If localist models of word identification are to remain serious competitors to distributed models, they must be designed on the basis of word recognition experiments and must incorporate learning algorithms that explain priming a by-product of learning. There is no reason to assume this cannot be done. Indeed, recent modeling based on the insights of Grossberg and colleagues (Grossberg, 1980; Carpenter & Grossberg, 1987) have shown how localist models that learn can support a number of long-term priming results (Davis, 1999).

Before concluding, it is worth considering a more general implication that follows from our finding that a learning process produces facilitatory and inhibitory priming effects for the rhyme and nonrhyme words, respectively. In a number of recent reports, it has been claimed that changes in sensitivity result in no costs. For example, Keane, Verfaellie, Gabrieli, and Wong (2000) wrote: “According to a sensitivity account, by contrast, only the benefit, and not the cost, should be observed: If priming improves the ability to extract perceptual information from a stimulus, then identification of a word should be enhanced by prior exposure to that word and should not be harmed by prior exposure to its orthographic mate” (p. 318). It has also been proposed that separate mechanisms underlie benefits and costs in priming (e.g., McKoon & Ratcliff, in press). Although the latter claim may prove to be correct, the present findings highlight the fact that benefits and costs can be the product of a learning process whose function is to improve word processing. A particularly compelling example of costs associated with learning can be found in a related language domain. At a few months, babies are able to perceive and produce the phonemes of all the languages of the world, but as they are exposed and learn the phonology of their particular language, they become less sensitive to key phonetic distinctions in other languages while becoming more adept in identifying the phonemes within their own language (Kuhl et al., 1992). For example, Japanese speakers have great difficulty in perceiving and producing the phonemes “l” and “r” in English. It seems unlikely that a separate bias mechanism underlies this deficit, just as there are no bias mechanisms underlying the costs we observed in the present simulations. Learning systems that improve processing do not rule out cost in performance, although the benefits outweigh costs in the domain in which they function.

Conclusion

The S&M(89) model which learns distributed word representations via back-propagation is able to accommodate the combination of robust long-term repetition and null form-priming effects as well as the interaction between frequency and priming reported in the literature. Indeed, the model made a novel prediction regarding the conditions in which form priming occurs, which was confirmed in a behavioral study. It appears that the slow interleaved learning associated with back-propagation can support priming following a single study episode.

APPENDIX A

Words and Pseudohomophones Used in Experiment 1

Target	Form related prime	Pseudohomophone
beek	beep	braik
boast	toast	braiv
brag	bran	brane
brash	trash	brude
bribe	tribe	cheet
chafe	chase	chuze
champ	chump	crait
chant	chart	dreem
cheap	cheat	faik
cheer	cheek	flore
chess	chest	fraze
clamp	clump	frute
cleft	cleat	gerl
cloud	clout	gleem
clove	glove	gloab
crab	crib	graid
creak	croak	graip
cube	cute	grait
deft	dent	grean
dunk	dusk	greef
fault	vault	grone
flake	flare	hert
fleck	flick	hite
flute	fluke	hoal
frown	drown	hoap
gorge	forge	howse
grand	grind	koil
guess	guest	leese
hinge	binge	leesh
lease	cease	lern
mall	malt	munny
mink	mint	munth
mule	mute	paist
prank	crank	peese
reek	reel	phoan
round	bound	reech
shark	stark	shure
shirt	skirt	smoak
shrug	shrub	soal
skull	skill	spaid
skunk	stunk	spaid
slab	slob	staik
sleet	sleek	stoan
smack	shack	sune
smoke	spoke	supe
snag	snug	teech
snout	scout	teeze
spoof	spoon	tode
stalk	stale	trane
steer	sneer	trupe
stilt	still	waik
swam	swan	werd
sworn	scorn	wheet
thumb	thump	whyte
trump	tramp	wreed
yearn	learn	wurld

APPENDIX B

Target and Prime Words Used in Experiment 2

Target	Form-related rhymes	Form-related nonrhymes
barn	darn	warn
bash	dash	wash
bear	pear	rear
blown	flown	clown
boot	hoot	soot
carp	harp	warp
cease	lease	tease
con	don	ton
cove	rove	dove
cross	dross	gross
crow	grow	brow
dart	cart	wart
dew	pew	sew
dorm	norm	worm
drown	frown	grown
fowl	howl	bowl
freak	creak	break
hood	wood	mood
hose	pose	dose
host	post	lost
lard	bard	ward
Mint	hint	pint
mow	tow	cow
pour	tour	dour
rough	tough	cough
sand	band	wand
tomb	womb	comb
tool	pool	wool

REFERENCES

Andrews, S. (1982). Phonological recoding: Is the regularity effect consistent? *Memory & Cognition*, **10,** 565–575.

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). *The CELEX lexical database (CD-ROM)*. Philadelphia, PA: Univ. of Pennsylvania, Linguistic Data Consortium.

Becker, S., Moscovitch, M., Behrmann, M., & Joordens, S. (1997). Long-term semantic priming: A computational account and empirical evidence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **23,** 1059–1082.

Besner, D. (1998). Basic processes in reading: Multiple routines in localist and connectionist models. In R. M. Klein & P. A. McMullen (Eds.), *Converging methods for understanding reading and dyslexia* (pp. 413–458). Cambridge, MA: MIT Press.

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.

Bowers, J. S. (1996). Different perceptual codes support priming for words and pseudowords: Was Morton right all along? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **22,** 1336–1353.

Bowers, J. S. (1999). Priming is not all bias: Commentary on Ratcliff and McKoon (1997). *Psychological Review*, **106,** 582–596.

Bowers, J. S. (1999). Challenging the widespread assumption that connectionism and distributed representations go hand-in-hand. Manuscript submitted for publication.

Bowers, J. S. (2000a). In defense of abstractionist theories of word identification and repetition priming. *Psychonomic Bulletin & Review*, **7,** 83–99.

Bowers, J. S. (2000b). The modality specific and non-specific components of long-term priming are frequency sensitive. *Memory & Cognition*, **28,** 406–414.

Bowers, J. S., & Michita, Y. (1998). An investigation into the structure and acquisition of orthographic knowledge: Evidence from cross-script Kanji-Hiragana priming. *Psychonomic Bulletin & Review*, **5,** 259–264.

Carpenter, G. A., & Grossberg, S. (1987). A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer Vision, Graphics, and Image Processing*, **37,** 54–115.

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud—Dual-route and parallel-distributed-processing approaches. *Psychological Review*, **100,** 589–608.

Davelaar, E., Coltheart, M., Besner, D., & Jonasson, J. T. (1978). Phonological recoding and lexical access. *Memory & Cognition*, **6,** 391–402.

Davis, C. (1999). *The self-organising lexical acquisition and recognition (SOLAR) model of visual word recognition*. Unpublished doctoral dissertation, University of New South Wales.

Forster, K. I. (1976). Accessing the mental lexicon. In R. J. Wales & E. Walker (Eds.), *New approaches to language mechanisms* (pp. 257–287). Amsterdam: North–Holland.

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.

Grainger, J., & Ferrand, L. (1996). Masked orthographic and phonological priming in visual word recognition and naming: Cross-task comparisons. *Journal of Memory and Language*, **35,** 623–647.

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review*, **103,** 518–565.

Grainger, J., & Jacobs, A. M. (1998). On localist connectionism and psychological science. In J. Grainger & A. M. Jacobs (Eds.), *Localist connectionist approaches to human cognition* (pp. 1–38). Mahwan, NJ: Erlbaum.

Grossberg, S. (1980). How does the brain build a cognitive code? *Psychological Review*, **87,** 1–51.

Grossberg, S. (1987). Competitive learning—From interactive activation to adaptive resonance. *Cognitive Science*, **11,** 23–63.

Grossberg, S., & Stone, G. (1986). Neural dynamics of word recognition and recall—Attentional priming, learning, and resonance. *Psychological Review*, **93,** 46–74.

Hanson, S. J., & Burr, D. J. (1990). What connectionist models learn: Learning and representation in connectionist networks. *Behavioral and Brain Sciences*, **13,** 471–518.

Hanson, V. L., & Wilkenfeld, D. (1985). Morphophonology and lexical organization in deaf readers. *Language and Speech*, **28,** 269–280.

Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review*, **106,** 491–528.

Keane, M. M., Verfaellie, M., Gabrieli, J. D. E., & Wong, B. M. (2000). Bias effects in perceptual identification: A neuropsychological investigation of the role of explicit memory. *Journal of Memory and Language*, **43,** 316–344.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experiences alter phonetic perception in infants by 6 months of age. *Science*, **255,** 606–608.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

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., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception. 1. An account of basic findings. *Psychological Review*, **88,** 375–407.

McClelland, J. L., & Rumelhart, D. E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General*, **114,** 159–188.

McClosky, M., & Cohen, N. J. (1989). Catastrophic interference in connectionist networks: The sequential learning problem. In G. H. Bowers (Ed.), *The psychology of learning and motivation* (pp. 109–165). New York: Academic Press.

McKoon, G., & Ratcliff, R. (in press). The counter model for word identification: Reply to Bowers (1999). *Psychological Review*.

Morton, J. (1979). Facilitation in word recognition: Experiments causing change in the logogen model. In P. A. Kolers, M. E. Wrolstad, & H. Bouma (Eds.), *Processing models of visible language* (pp. 259–268). New York: Plenum.

Murrell, G. A., & Morton, J. (1974). Word recognition and morphemic structure. *Journal of Experimental Psychology*, **102,** 963–968.

Napps, S. E., & Fowler, C. A. (1987). Formal relationships among words and the organization of the mental lexicon. *Journal of Psycholinguistic Research*, **16,** 257–272.

Norris, D. (1994). A quantitative multiple-levels model of reading aloud. *Journal of Experimental Psychology: Human Perception and Performance*, **20,** 1212–1232.

Page, M. (2000). Connectionist modeling in psychology: A localist manifesto. *Behavioral and Brain Sciences*, **23,** 443–512.

Peereman, R., Content, A., & Bonin, P. (1998). Is perception a two-way street? The case of feedback consistency in visual word recognition. *Journal of Memory and Language*, **39,** 151–174.

Pexman, P. M., Lupker, S. J., Jared, D., Toplak, M., & Rouibah, A. (1996). *Strategically controlling phonology in response to pseudohomophones?* Paper presented at the 37th annual meeting of the Psychonomic Society, Chicago, IL.

Plaut, D. C. (1997). Structure and function in the lexical system: Insights from distributed models of word reading and lexical decision. *Language and Cognitive Processes*, **12,** 765–805.

Plaut, D. C., & Booth, J. R. (2000). Individual and developmental differences in semantic priming: Empirical and computational support for a single-mechanism account of lexical processing. *Psychological Review*, **107,** 786–823.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, **103,** 56–115.

Rapp, B. C. (1992). The nature of sublexical orthographic organization: The bigram trough hypothesis examined. *Journal of Memory and Language*, **31,** 33–53.

Ratcliff, R. (1990). Connectionist models of recognition memory: Constraints imposed by learning and forgetting functions. *Psychological Review*, **97,** 285–308.

Ratcliff, R., & McKoon, G. (1997). A counter model for implicit priming in perceptual word identification. *Psychological Review*, **104,** 319–343.

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., & Mathew, S. (1999). Implicit memory for phonological processes in visual stem completion. *Memory & Cognition*, **27,** 1–11.

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.

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

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

Squire, L. R., Shimamura, A. A., & Graf, P. (1987). Strength and duration of priming effects in normal subjects and amnesic subjects. *Neuropsychologia*, **25,** 195–210.

Stark, C. E. L., & McClelland, J. L. (2000). Repetition priming of words, pseudowords, and nonwords. *Journal of Experimental Psychology: Learning Memory and Cognition*, **26,** 945–972.

Stone, G. O., Vanhoy, M., & VanOrden, G. C. (1997). Perception is a two-way street: Feedforward and feedback phonology in visual word recognition. *Journal of Memory and Language*, **36,** 337–359.

Stone, G. O., & VanOrden, G. C. (1993). Strategic control of processing in word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, **19,** 744–774.

Subramaniam, S., Biederman, I., & Madigan, S. (2000). Accurate identification but no priming and chance recognition memory for pictures in RSVP sequences. *Visual Cognition*, **7,** 511–535.

Tovee, J. J., & Rolls, E. T. (1995). Information encoding in short firing rate epochs by single neurons in the primate temporal visual cortex. *Visual Cognition*, **2,** 35–59.

Treiman, R., Mullennix, J., Bijeljac-Babic, R., & Richmond-Welty, E. D. (1995). The special role of rimes in the description, use, and acquisition of English orthography. *Journal of Experimental Psychology: General*, **124,** 107–136.

Wagenmakers, E. M., Zeelenberg, R., & Raaijmakers, J. G. W. (2000). Testing the counter model for perceptual identification: Effects of repetition priming and word frequency. *Psychonomic Bulletin & Review*, **7,** 718–726.

(Received February 22, 2000)
(Revision received February 7, 2001)