THEORETICAL NOTES

Counter Model for Word Identification: Reply to Bowers (1999)

Gail McKoon and Roger Ratcliff
Northwestern University

The counter model (R. Ratcliff & G. McKoon, 1997) was designed to explain the normal processes of word identification and how they are influenced by a prior encounter with a word. The model accounts for the findings of word identification experiments in which words are flashed briefly. A crucial finding is that prior encounters with words typically lead to biases such that a previously encountered word is more likely to be given as a response. However, for low-frequency words, a prior encounter can improve overall performance (J. S. Bowers, 1999; E. M. Wagenmakers, R. Zeelenberg, & J. G. W. Raaijmakers, 2000). The authors show how the model can explain this result. Also, J. S. Bowers (1999) has claimed that some earlier data concerning dissimilar alternatives in forced-choice experiments that support the counter model are spurious, but the authors show that his claims are incorrect. In sum, the authors argue for a theoretical approach that offers a detailed description of the cognitive processes of word identification and explains performance across tasks, measures, and independent variables.

For the past 15 years, the widely ranging phenomena that fall under the umbrella label “implicit memory” have intrigued researchers. Many verbal hypotheses have been put forward to describe one or another of these phenomena, but detailed theoretical models have appeared only recently (Ratcliff & McKoon, 1997; Schooler, Shiffrin, & Raaijmakers, 2001). In this article, we present new evidence to support and extend one of these models, the counter model for word identification. Also, we show that empirical results that have been thought to contradict the model can in fact be explained by mechanisms already built into the model. In so doing, we demonstrate the strength of a theoretical approach that allows an explicit account of the cognitive processes that underlie performance.

The goal in proposing the counter model (Ratcliff & McKoon, 1997) is to embed an explanation of implicit effects, specifically, priming in masked word identification, into an information processing model that is designed to perform the main functions of word identification (see also Ratcliff & McKoon, 1996). Priming is implemented as a by-product of those functions (Morton, 1970). The aim is to account both qualitatively and quantitatively for multiple empirical effects from multiple tasks. The accumulated data provide severe constraints on possible models. The data led us to a model for which the central assumption about priming is that prior encounters with a word bias the decision process that interprets perceptual information. This assumption contradicts what many previous researchers have postulated, namely, that prior encounters affect the representation of the word itself in memory in such a way that processing on subsequent encounters is facilitated.

The counter model is a variant of Morton’s (1969, 1970, 1979) logogen model for word identification. Each word in memory is assumed to have a counter. These counters are decision counters in that the accumulation of counts in them is the decision mechanism by which the system identifies a word that is presented to it. When a word is presented, at each unit of time one count is accumulated to one and only one counter. Under normal reading conditions, most of the counts are controlled by perceptual features of the stimulus; that is, they will be accumulated in counters that are appropriate for the stimulus. For example, a count from a perceptual feature of the letter d would be accumulated by the counter for some word that contained a d. However, in perceptual word identification experiments, a word is flashed for only a very short time and then masked. Under these impoverished stimulus conditions, many counts are not determined by the stimulus; these null counts represent noise in the system and can be accumulated by any counter. The accumulation of counts of random noise allows the system to make a decision and respond even when there is little or no perceptual information from the stimulus. The system reaches a decision when the number of counts in one counter exceeds the number of counts in the maximum of the other counters by some criterion amount.

To explain priming, a prior encounter with a word is assumed to result in the word’s decision counter attracting more counts than it otherwise would. The counters in the system are assumed to be arranged in terms of similarity, with similar words close together in similarity cohorts. The attractive force attributable to prior exposure is assumed to have only a short range, and so a counter can only attract counts away from other counters in its cohort of

Gail McKoon and Roger Ratcliff, Department of Psychology, Northwestern University.

Preparation of this article was supported by National Institute on Deafness and Other Communication Disorders Grant R01-DC01240 and National Institute of Mental Health Grant HD MH44640. We thank Michael Masson and Jeroen Raaijmakers for sharing their experimental results with us.

Correspondence concerning this article should be addressed to Gail McKoon, Department of Psychology, Northwestern University, Evanston, Illinois 60208.
similar neighbors; it cannot attract counts away from counters of dissimilar words.

The counter model is a departure from and a challenge to previous ideas about priming in implicit tasks. The critical difference is that the counter model predicts that the effect of prior encounters is typically one of bias, not facilitation, that is, not an increase in the ease with which a previously encountered word is perceived. In the counter model, a prior encounter biases the decision process to respond with the previously encountered word, whether or not it is the correct response. According to other views, what a prior encounter does is facilitate the processing of the correct response, for example, by changing the resting level of activation for the word or by creating a new representation of the word in a word form memory system (Schacter & Tulving, 1994; Tulving & Schacter, 1990).

With the assumptions just outlined, the counter model explains a number of findings in perceptual identification experiments, including findings with a variety of tasks, independent variables, and dependent variables (e.g., Masson, 1998; Masson & MacLeod, 1996; Raaijmakers, Schooler, & Shiffrin, 1997; Ratcliff, Allbritton, & McKoon, 1997; Ratcliff & McKoon, 1997; Ratcliff, McKoon, & Verwey, 1989). First, the model accounts for the effects of two standard word identification variables, word frequency and cohort size, in both forced-choice and naming tasks. Second, the model accounts for the bias effect that results from prior encounters with a word in forced-choice and naming experiments. Third, the model explains why bias occurs in forced choice only when the two alternative choices are similar to each other, not dissimilar. Fourth, the model accounts for the amount of bias as a function of stimulus duration, with only one parameter varying, and correctly predicts that bias should occur even when the amount of information from the stimulus is so small that performance is near chance. Fifth, the model fits the data from a “yes/no” identification version of perceptual identification. Sixth, with only one parameter varying, the model fits the data from individual participants simultaneously for the naming and forced-choice tasks. Finally, unlike Morton’s (1969, 1970) original model, the counter model appropriately decouples the effects of word frequency and priming.

Recently, Bowers (1999) has put forward a defense of the more traditional views of priming in perceptual processing, including a series of eight experiments from which the results, he claims, contradict the counter model. Here we refute each of his claims. In the course of refuting Bowers’ claims, we reiterate the strength of the counter model in providing a quantitatively explicit account of a wide range of perceptual identification data. We discuss and document new empirical findings about the perception of low-frequency words (Bowers, 1999; Wagenmakers, Zeelenberg, & Raaijmakers, 2000) and show how the model is capable of explaining these effects while still retaining its power to explain all the original data. Although the explanation of the new findings means a change in our understanding of one of the parameters of the model, there is no change at all in the structure of the model.

Bias

The traditional view of priming in word identification is that a prior encounter with a word leads to the facilitation of processing on subsequent encounters, perhaps because a new representation of the word is encoded in memory or because the resting level of activation of the word is raised. In contrast, in the counter model, the effect of a prior encounter is not usually facilitation of processing. Instead, it is a bias in the decision process: The probability of correctly identifying a word is increased by a prior encounter with the word itself but it is decreased by a prior encounter with a similar but different word.

Bias appears with both naming and forced-choice perceptual identification procedures (Ratcliff & McKoon, 1997; Ratcliff et al., 1989). In a typical experiment, a list of words is presented to participants, and these words plus others are subsequently tested in perceptual identification. With the naming perceptual identification procedure, a target word is flashed and then masked, and the participant’s task is to name the target word aloud. Prior presentation of a target word increases the probability of a correct response for that word, and prior presentation of a word similar to the target increases the probability of an error, that is, responding with the similar word instead of the target. This bias effect occurs in the counter model because the counter for a previously presented word attracts counts away from the counters of its similar neighbors.

With the forced-choice procedure, a target word is flashed briefly and then masked, and then two alternative words are presented. Participants are asked to decide which of the alternatives matches the word that was flashed. Again, the data show bias: If died was the flashed target word and it was previously presented, then there is an increase in the probability of correctly choosing died from the similar alternatives died and lied. However, if lived was previously presented, there is a decrease in the probability of correctly choosing died from the alternatives died and lied. In other words, for the target died, there is a benefit from previous study of died and a cost from previous study of lied. However, this bias effect appears only when the two alternatives are similar to each other, not when they are dissimilar (e.g., died and sofa). With dissimilar alternatives, neither we nor Masson found any significant effect of prior study: no benefit from prior presentation of the target word and no cost from prior presentation of the alternative (Masson, 1998; Ratcliff et al., 1977; Ratcliff & McKoon, 1997; Ratcliff et al., 1989; Rouder, Ratcliff, & McKoon, 2000).

Bias in forced choice is explained by the counter model with the same mechanism as that used for naming: The counter for a previously encountered word attracts counts away from its neighbors. There is bias for similar but not dissimilar alternatives because the attraction force extends only through the cohort of a word’s similar neighbors.

Bowers (1999) singles out for attack the finding of bias in forced choice with similar but not dissimilar alternatives, not mentioning bias in naming, and he argues that the failure to find bias with dissimilar alternatives is spurious. In four experiments, he finds bias with both similar and dissimilar alternatives. This finding leads him to reject the counter model.

In response, we call attention to a crucial difference between the forced-choice procedure used by Bowers (1999) and the procedures used by others. Bowers’ experiments were described to the participants as experiments about “word perception,” and the participants were given no instructions about the lists of words that were presented prior to the perceptual identification tests, except to read them. In the procedure that we use (like that used by Masson, 1998), the lists are called “study lists,” and participants are instructed to memorize the words in the lists for a later memory test.
Our guess was that participants in Bowers' experiments thought that the lists were presented in order to help them on the perceptual identification tests because no purpose was given for the lists of words that preceded the perceptual identification tests. Consequently, on some proportion of the trials, they adopted a strategy of choosing whichever of the forced-choice alternatives had been on the previously presented list, both when the alternatives were similar and when they were dissimilar. With the instructions that we use, participants are encouraged not to think that the lists of words are presented to help them on the perceptual identification tests, because they are given a different purpose for the lists of words, namely, to memorize them; therefore, the participants in our experiments would not consciously tend to select alternatives from the study lists. We thought that this difference in strategies, brought about by the difference in instructions, might have led to the difference between Bowers' results and those obtained by Masson and by us, and we tested this idea in Experiment 1.

**Experiment 1**

The aim of Experiment 1 was to show that bias is obtained in forced choice with dissimilar alternatives only with Bowers' (1999) instructions, not with the instructions that have been standardized. The standard instructions emphasize memorizing for a later memory test the lists of words presented prior to the perceptual identification tests. With Bowers' instructions—simply to read the lists of words—we expected participants to try to use the words from the lists in making their forced-choice decisions (producing a bias effect with dissimilar alternatives) and also to report after the experiment that they had done so.

**Method**

The experiment was modeled on Bowers' (1999) Experiments 1, 2, and 3. The experiment consisted of four blocks, each block composed of a list of 12 words (presented for 1 s each) followed by 24 perceptual identification test items. The four blocks were preceded by 24 practice perceptual identification tests. The pairs of words used in the experiment were randomly chosen from 168 pairs used by Ratcliff and McKoon (1997), the two words of a pair as dissimilar as possible in the shapes of their letters (e.g., died-sofa). The pairs were assigned randomly to conditions in the experiment.

For the perceptual identification tests, there were three conditions (see Table 1): Either the flashed target word or its dissimilar pair mate was presented in the preceding list or neither word was presented.

For 14 participants, the instructions were like those that we have used in earlier experiments (Ratcliff et al., 1997, Experiment 3; Ratcliff & McKoon, 1997, Experiments 1 and 5; Ratcliff et al., 1989, Experiment 5), in which we found no bias with dissimilar alternatives. The lists of words were called study lists, and the participants were instructed to memorize the words in the lists for a memory test to be given at the end of the experiment. At the beginning of each study list, they were reminded to try to memorize the words. For another 14 participants, the instructions were like those used by Bowers (1999). The participants were told nothing about the lists of words, except to read them carefully, and they were reminded to read carefully at the beginning of each list. All 28 participants were told that the perceptual identification task was very difficult and that they should concentrate carefully to do their best. The participants were Northwestern University undergraduates participating in the experiment for credit in an introductory psychology class.

The stimulus materials were displayed on a gray-scale PC monitor with a refresh rate of 4 ms. This means that the beam on the computer monitor was controlled to return to the top of the screen after moving down one-fourth of the normal distance to the bottom. This procedure resulted in an effective screen size one-fourth the regular size but with a refresh rate of 4 ms instead of 16.67 ms (Von Brisinger, 1994). Responses were collected by use of the PC keyboard.

The perceptual identification stimuli were displayed as follows: a row of minus signs for 400 ms as a warning signal, a blank screen for 300 ms, the target word flashed for 20 ms, a mask (made up of random lines) displayed for 300 ms and covering an area slightly larger than that of the target, and the two forced-choice alternatives, presented side by side. Participants were instructed to press the 7/ key or the Z key to indicate which of the alternatives had been flashed. Which response was correct, the right alternative or the left one, was chosen randomly. After the response, the warning signal began the sequence for the next test.

**Results and Discussion**

After the experiment, we asked each participant whether he or she had used the word lists to help on the word identification tests. Although in general participants cannot be assumed to have good insights into their cognitive processes (see McKoon & Ratcliff, 1996, for a discussion of this issue), it is interesting to note here what participants said. None of the participants receiving the standard memorization instructions said that they had intentionally chosen words from the study lists on the forced-choice tests. In contrast, all of the participants receiving Bowers' (1999) instructions either volunteered or agreed that they had sometimes done so.\(^1\)

---

1 Bowers (1999) attempted to rule out the possibility that explicit retrieval played a role in his experiments by using a deadline procedure (Bowers, 1999, Experiment 5). Participants were instructed to make their forced-choice decisions exactly 750 ms after the two alternatives were displayed. The results of this experiment were supposed to show that bias was obtained with dissimilar alternatives even when a deadline was used to eliminate explicit retrieval. There are two problems. One is that bias was not obtained; with dissimilar alternatives, there was an increase in the probability of a correct answer both when the target was previously presented and when the dissimilar alternative was previously presented. Second, the time course for the retrieval of information from explicit memory for words is very rapid; data have shown that such information becomes available after only 350 ms (Dosher, McElree, Hood, & Rosedale, 1989; Gronlund & Ratcliff, 1989; Wickelgren, Corbett, & Dosher, 1980). A 750-ms deadline therefore cannot be said to eliminate explicit retrieval for words.

**Table 1**

<table>
<thead>
<tr>
<th>Experiment 1 Data: Mean Probability of a Correct Response in Forced Choice With Dissimilar Alternatives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability under the following test conditions (example: flashed target—died; forced-choice alternatives—died and sofa):</td>
</tr>
<tr>
<td>Instructions</td>
</tr>
<tr>
<td>Standard (memorize)</td>
</tr>
<tr>
<td>Bowers's</td>
</tr>
</tbody>
</table>

---

\(^{1}\)
Table 1 shows the mean probability of a correct response for each condition. Bowers (1999) suggested that performance has been near ceiling in some perceptual identification experiments, but that is not the case here; the probability of a correct response in the baseline condition is only .61.

With the memorization instructions, the results replicate the usual finding—no significant effect of prior study. However, with Bowers’ (1999) instructions, the results show bias: an increased probability of a correct response over baseline when the flashed target was presented in the preceding list of words and a decreased probability of a correct response when the alternative was presented in the preceding list. In other words, if the two alternatives were died and sofa and the flashed target was died and if died had appeared on the previous lists of words, participants receiving Bowers’ instructions were more likely to choose died; if sofa had appeared on the previous lists of words, they were more likely to choose sofa. Analysis of variance (with test condition and instruction condition as variables) showed a significant interaction, $F(2, 52) = 3.4, p < .05$ (statistical significance was set at $p < .05$ throughout this study), reflecting the finding of bias with Bowers’ instructions but not with standard instructions. The standard error of the mean for the probability of a correct response was .02.

We conclude that Bowers’ (1999) experiments showed bias in forced choice with dissimilar alternatives because of the instructions that were used, instructions that suggested to participants that they should use in their forced-choice decisions words from the lists that preceded them. Given another purpose for the lists of words (i.e., to memorize them for a later test), bias disappears. Therefore, Bowers’ finding of bias with dissimilar alternatives is not a reason to reject the counter model.

Bias With No Perceptual Information From the Stimulus

One of the consequences of random noise (null counts) in the counter model is that the system can make a response even when perceptual information is largely or even completely absent, which is something that subjects do in perceptual identification experiments. Moreover, no matter how few of the counts are controlled by perceptual information, the remaining (null) counts are still subject to the attraction caused by prior study, and so performance can still show bias. In forced choice, when the two alternatives are presented, the accumulation of counts is restricted to the counters for the two alternatives. Even when the accumulation of counts is determined mostly by noise, one of the counters will eventually accumulate more counts than the other by the criterion amount, and so a response can be made. In addition, if one of the alternatives was previously studied, it will be more likely to be the “winner” because of theft of null counts from the other alternative (as long as the two alternatives are similar). Ratcliff & McKoon (1997, Experiment 4) confirmed this prediction of the counter model. We used a flash time so short that baseline performance (previous study of neither alternative) was at chance, meaning that most counts must have been null counts; we found that forced choice with similar alternatives still showed the bias effect.

Bowers (1999) also found a bias effect with similar alternatives when performance was at chance. In his experiment, performance was at chance not because of a very short flash time for the target word but because only strings of random symbols were flashed—there was no target word at all. Bowers suggests that his finding is inconsistent with the counter model but, in fact, as just reviewed, it is exactly predicted by the model.

Bias Versus Facilitation

For forced choice, prior study of a target word leads to a benefit to performance, and prior study of a similar word leads to a cost. The benefit and the cost are typically of about the same size, and so there is no overall facilitation in performance. This finding of no overall improvement in performance is surprising from the point of view of traditional ideas about implicit memory, and so Ratcliff and McKoon (1997, Experiment 2) provided converging evidence for it. The idea was that an overall improvement might be found if there was prior study of both alternatives. If, for example, the flashed test word was died and the forced-choice alternatives were died and lied, then both died and lied were presented in the study list. However, there was no significant effect of prior study at all. According to the counter model, the prior encounters with the words increased the attractive power of both their counters equally, canceling each other out.

However, Bowers (1999) and Wagenmakers et al. (2000) have recently found that prior study leads to a true improvement in performance for low-frequency words. We verified this result in Experiment 2, described below.²

In the counter model, a true improvement in performance occurs when there is better extraction of information from the stimulus. The parameter of the model that captures this notion is $ps$, the probability that a count is based on perceptual information, that is, the probability that it is not null. In past applications, $ps$ varied with the flash time of the target. The longer the flash time, the greater the value of $ps$.

The counter model can account for an improvement in performance for low-frequency words by assuming that prior study affects the parameter $ps$. There are two possibilities. One is that without prior study, low-frequency words have a lower value of $ps$ than high-frequency words. The idea is that low-frequency words are unfamiliar, and so their features (e.g., letter combinations) are more difficult or time-consuming to extract. With prior study, the value of $ps$ increases. The second possibility is that unfamiliarity with low-frequency words presented as forced-choice alternatives can sometimes, on some proportion of trials, interrupt processing of the flashed target so that no information from the target is available; in this situation, participants would have to guess which of the two choices was the correct one.

The key point is that in the counter model, an improvement in performance is attributed to a mechanism other than bias. To demonstrate this point, Experiments 3 and 4 dissociated the improvement in performance for low-frequency words from bias by using a 48-hr delay interval between study and test. A delay of this length is sufficient for bias (or priming) to dissipate (for bias or

² In earlier experiments, we found no true improvement in performance for either low- or high-frequency words (Ratcliff & McKoon, 1997). However, the relevant results were post hoc partitions of data, with data combined across manipulations of a number of other experimental variables, including, for example, whether type font was the same or different at study versus test. We attribute this earlier failure to find an improvement for low-frequency words to a lack of power and the presence of the other manipulations.
priming effects of the size at issue here) (McBride & Dosher, 1997; Ratcliff & McKoon, 1997, Experiment 7; Whitlow & Dalton, 1999; see also Feustel, Shiffrin, & Salasoo, 1983). In contrast, we expected that the improvement in processing for low-frequency words might survive the delay.

We first discuss Experiment 2, which verified the improvement in performance for low-frequency but not high-frequency words that was found by Bowers (1999) and by Wagenmakers et al. (2000). Then, in Experiments 3 and 4, we introduce the 48-hr delay.

**Experiment 2**

We used a forced-choice procedure with our standard instructions to memorize the study lists of words. The frequencies of the words were varied (see Table 2). There were two test conditions: Both the flashed target word and the alternative word were presented in the previously studied list of words or neither the target nor the alternative was studied.

For high-frequency words, perceptual information should be extracted efficiently, and prior study should not change that. The only effect of prior study of both alternatives should be an increase in the attractive power of both, and so there should be no change in the probability of a correct response. For low-frequency words, prior study should increase the extraction of perceptual information, leading to a genuine improvement in performance. Prior study should also increase the attractive power of both alternatives but, as for high-frequency words, the increase in attractive power should have no effect on performance.

**Method**

There were three pools of words: high frequency, 60 pairs with both words having Kucera-Francis (Kucera & Francis, 1967) frequencies of greater than 78; low frequency, 60 pairs with both words having a Kucera-Francis frequency of 4 or 5; and very low frequency, 60 pairs with a Kucera-Francis frequency of 1 or 0. The words of a pair were as dissimilar as possible in the shapes of their letters. We used dissimilar pairs only, not similar pairs, to reduce the number of conditions in the experiment, thereby increasing power. The numbers of letters per word ranged from four to seven. The words were assigned randomly to the two test conditions—either the target word and its pair mate were studied or neither the target nor its pair mate was studied. The procedure was the same as that used in Experiment 1. There were five blocks, each block with a list of 18 words (both words of 9 pairs) to study and 18 forced-choice test items. The target words were flashed for 20 ms and masked in the same way as in Experiment 1. Forty participants were selected from the same population as that used in Experiment 1.

**Results**

Table 2 shows the mean probability of a correct response for each condition. Baseline performance is approximately in the middle of the possible range (with .50 being chance and 1.00 being perfect performance).

For the high-frequency words, the usual result was obtained—no effect of prior study. However, for the low- and very-low-frequency words (which were combined for statistical tests), prior study of the alternatives improved performance. The interaction between test condition and word frequency was significant, $F(1, 39) = 4.5$. Planned tests showed that without prior study, performance for the low- and very-low-frequency words was worse than performance for the high-frequency words, $F(1, 40) = 4.5$, but with prior study, performance for the low- and very-low-frequency words was as good as (and not significantly different from) performance for the high-frequency words, $F(1, 40) = 2.0$. The standard error of the mean for the probability of a correct response was .01.

**Experiments 3 and 4**

The improvement in performance demonstrated in Experiment 2 for low-frequency words confirms findings obtained by Bowers (1999) and Wagenmakers et al. (2000). For the counter model, we interpret this improvement to be the result of either an increase in the quality of the perceptual information extracted from the flashed target or a decrease in the probability that low-frequency alternatives derail the processing of the flashed target. Either way, we thought that the improvement in performance might be a relatively long-lasting change, and we tested this idea in Experiments 3 and 4 by interposing a 2-day delay between the time when target words were studied and the time when they were tested. On the other hand, we expected bias to be relatively short-lived.

**Method**

The participants each took part in two sessions separated by approximately 48 hr. In the first session, words were presented for study with a study–test recognition memory procedure. In the second session, the procedure was much the same as that used in Experiments 1 and 2, with blocks of words to study and with each block followed by perceptual identification tests.

In Experiment 3, only dissimilar pairs of very-low-frequency words were used, the same pool as that used in Experiment 2. Both words of a pair were studied in both sessions, the first session only, the second session only, or neither session (see Table 3). There were 15 pairs of words (chosen randomly) for each condition for each participant.

In Experiment 4, the very-low-frequency, dissimilar pairs of words were used, as was a set of pairs of similar, higher frequency words. The higher frequency words were chosen from the pool of words used in Experiment 1 such that they all had frequencies of greater than 5. For the low-frequency pairs, both words of a pair were studied either in the first session or in neither session. There were 14 pairs of words for each participant for each condition. For the high-frequency pairs, either the flashed target or its similar alternative was studied in the first session or in the second session.

---

**Table 2**

**Experiment 2 Data: Mean Probability of a Correct Response in Forced Choice With Dissimilar Alternatives**

<table>
<thead>
<tr>
<th>Target word (died) and alternative (sofa) presented in preceding list of words</th>
<th>Neither alternative presented in preceding list of words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability under the following test conditions (example: flashed target—died; forced-choice alternatives—died and sofa):</td>
<td></td>
</tr>
<tr>
<td>Words</td>
<td>Very low frequency</td>
</tr>
<tr>
<td></td>
<td>Low frequency</td>
</tr>
<tr>
<td></td>
<td>High frequency</td>
</tr>
</tbody>
</table>
THEORETICAL NOTES

Table 3

Experiment 3 and 4 Data: Mean Probability of a Correct Response in Forced Choice With Dissimilar Alternatives

<table>
<thead>
<tr>
<th>Study session</th>
<th>Experiment</th>
<th>Target and alternative</th>
<th>Neither</th>
<th>Target</th>
<th>Alternative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low-frequency words</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First and second</td>
<td>3</td>
<td>.72</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First</td>
<td>3</td>
<td>.72</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Second</td>
<td>3</td>
<td>.73</td>
<td></td>
<td>.67</td>
<td>.79</td>
</tr>
<tr>
<td>Neither</td>
<td>3</td>
<td></td>
<td>.79</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High-frequency words</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First</td>
<td>4</td>
<td>.79</td>
<td>.77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Second</td>
<td>4</td>
<td>.82</td>
<td>.73</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(see Table 3). There were 24 pairs of words for each participant for each condition.

In the first session of Experiment 3, participants studied two lists of 30 words per list (1 s per word), each list followed by 60 test items, the 30 words from the study list plus 30 new words. In the first session of Experiment 4, participants studied two lists of 38 words per list, each list followed by 70 test items, the 38 words from the study list plus 32 new words. In both experiments, participants were instructed to study the words in the lists and respond to each test word by pressing the "F" key if the word was in the study list and the "Z" key if it was not. After the two lists of words were presented, the same words were presented for study again in two more lists, in a different random order and with different lures in the test lists.

For the second sessions of both experiments, the procedure was the same as that used in Experiment 2, with the standard instructions to memorize the lists of words. For Experiment 3, there were three blocks with 20 words studied and 20 words tested for each block. For Experiment 4, there were two blocks with 24 words studied in each block (12 high-frequency target words and 12 high-frequency alternatives; see Table 3) and 62 words tested for each block (14 low-frequency targets and 48 high-frequency targets). The times for which target words were flashed were chosen for each participant on the basis of his or her performance on practice trials. For Experiment 3, each participant's flash time was 10, 12, or 14 ms, and for Experiment 4, each participant's flash time was 12, 14, or 16 ms. There were 40 participants for Experiment 3 and 34 for Experiment 4, selected from the same population as that used in Experiment 1.

Results and Discussion

The mean probabilities of a correct response for the data from the second sessions are shown in Table 3. Again, baseline performance is in the middle of the range between chance and ceiling. Recognition performance in the first session was good: For Experiment 3, 89% correct for old items and 90% correct for new items (a d' value of 2.5); for Experiment 4, 87% correct for old items and 85% correct for new items (a d' value of 2.2).

For low-frequency words in Experiments 3 and 4, study of both members of a pair prior to the perceptual identification tests improved the probability of a correct response by about .06 over baseline (which was .67 in Experiment 3 and .79 in Experiment 4). The increase in performance was about the same whether study occurred immediately prior to the perceptual identification tests, 48 hr earlier, or both. As predicted, the benefit of prior study lasted across the 48-hr delay. For Experiment 3, the interaction between whether or not a pair was studied in the first session and whether or not it was studied in the second session was significant, F(1, 39) = 6.5, indicating that performance was worse for words never studied than it was for any of the other conditions. The standard error of the mean for the probability of a correct response was .01. For Experiment 4, performance when a pair was studied in the first session was significantly better than that when a pair was never studied, F(1, 33) = 5.4. The standard error of the mean for the probability of a correct response was .02.

In contrast to the benefit for the low-frequency words, bias for high-frequency words was observed only when the words were studied immediately prior to the perceptual identification tests. There was no significant bias toward a studied word after the 48-hr delay. The interaction between whether the target or its alternative was studied and whether study occurred in the first or second session was significant, F(1, 33) = 5.0. The standard error of the mean for the probability of a correct response was .01.

We conclude that the improvement in performance for low-frequency words, lasting over 2 days, is not the same as bias, which dissipated in the same time period. In the counter model, the improvement for low-frequency words is attributable either to an increase in the proportion of counts that are based on perceptual information or to a decrease in the proportion of trials for which low-frequency alternatives derail processing of the target. Either of these is quite separate from the attraction mechanism that explains bias.

Conclusions for Experiments 1, 2, 3, and 4

Bowers (1999) criticized the counter model in three main ways. First, he argued that bias appears in forced choice not just with similar alternatives but also with dissimilar alternatives, in contradiction to the model. In response, Experiment 1 shows that his finding of bias with dissimilar alternatives is attributable to instructions that suggest to participants that they should use the lists of words presented prior to the forced-choice tests to help them make their choices.

Second, he argued that the counter model is contradicted by bias that occurs in forced choice even when the flashed target is a string
of random symbols, not a word. However, this is exactly what the model predicts.

Finally, Bowers (1999) (as well as Wagenmakers et al., 2000) suggested that the model is falsified by the result that low-frequency words show an improvement in performance as a function of prior study. As just discussed for Experiments 2, 3, and 4, this result does not falsify the model. Given the available data when the counter model was developed, the only variable we knew to affect \( ps \) in the model was the flash time of the target word. However, it now appears that \( ps \) for low-frequency words can be affected by prior study. With this simple change, the model fits Bowers’ and Wagenmakers et al.’s (2000) new data well (see also Ratcliff & McKoon, 2000). It is important to stress that the model itself has not changed; only the assumption about how the parameter \( ps \) behaves has changed. The model still has all of the same basic mechanisms as when it was proposed in 1997: Counts are accumulated to decision counters for each unit of time, some counts are perceptually based and others are noise, counters are arranged in neighborhoods defined by similarity, the counters of high-frequency words have higher resting levels than the counters of low-frequency words, prior study allows a counter to attract more counts than it otherwise would, the proportion of counts that are based on perception depends on the quality of the information extracted from the stimulus, and a decision is made when the counter for one word exceeds the maximum of the counters for all the other words in the system. These mechanisms make the counter model a fully specified model that provides quantitative predictions about performance across a range of tasks and variables.

Bowers’ Alternative Interpretations of Priming in Perceptual Word Identification

Bowers (1999, pp. 586–587) offers several alternative interpretations of the effects of prior encounters with a word on perceptual identification of the word. Each one is offered as a possible account of some part of the available data, for example, the effect of one independent variable on performance in one kind of task. Bowers suggests the following: a prior encounter with a word leads to an improvement in perceptual processing, an “increased sensitivity”; in forced choice, “the test alternatives themselves are primed” such that any improvement in performance attributable to a prior encounter is canceled out by priming of the alternatives; there is no improvement in performance in forced choice with dissimilar alternatives because decisions about dissimilar alternatives are made on a different basis than decisions about similar alternatives; and “priming for high-frequency words is largely mediated by bias” but “priming for low-frequency words is mediated, in part, by an increased sensitivity.”

These alternative interpretations can be refuted. For one thing, they are directed only at data from forced choice and do not account for data from naming or “yes/no” identification (Ratcliff & McKoon, 1997). Furthermore, the mechanism suggested to explain why there is no effect of prior study on forced choice with dissimilar alternatives cannot work. Without this mechanism, there should be, according to Bowers (1999), bias both when the test alternatives are similar and when they are dissimilar. However, bias with dissimilar alternatives appears only with Bowers’ instructions, not the standard instructions (Masson, 1998; Ratcliff et al., 1997; Ratcliff & McKoon, 1997; Ratcliff et al., 1989).

Bowers (1999, p. 592) concludes that “there is still a plausible and parsimonious explanation” for the effects of prior study in perceptual word identification “that is consistent with traditional views of priming.” However, the traditional views do not provide a coherent theory of word identification processes as a whole. There is no explicit account of the normal processes of word identification, there are no quantitative accounts of word identification or priming data, and hypotheses are not worked out in sufficient specificity to provide a comprehensive and explanation of experimental data.

The structure and mechanisms of the counter model establish restrictions on which mechanisms apply under which empirical conditions, and the structure and mechanisms define both qualitative and quantitative patterns of data. Without the restrictions provided by an explicit and tightly formulated theory, we cannot predict new findings about word identification, and so we cannot believe that we have even begun to understand how words are identified. To make progress, more models of the integrative sort represented by the counter model are needed. With the recent appearance of another such model (Schooler, Shiffrin, & Raaijmakers, 2001), the stage is set for competitive model testing.

General Conclusions

The experiments reported by Bowers (1999) and Wagenmakers et al. (2000) have helped to hone the experimental methods used to assess priming in word identification, and they have also helped to illustrate the power of detailed, quantitative modeling. For methodology, it is clear that instructions given to participants require some care, with experimenters being fully briefed on what to say and how to say it, and that participants should be debriefed to ascertain what they think are the purposes of various parts of experimental procedures. For modeling, the effect of prior study of low-frequency words is accommodated by the counter model with

---

3 The question might arise of how the counter model explains the effect of prior study of low-frequency words in naming. In typical data (e.g., Jacoby & Dallas, 1981, Experiment 3), the advantage of prior study for low-frequency words is greater than the advantage for high-frequency words. The probability of correctly naming a flashed target without prior study might be .6 for high-frequency words and .3 for low-frequency words, increasing to .8 and .6, respectively, with prior study. In earlier applications of the counter model, the model did accommodate a larger effect of prior study for low-frequency words, but the effect predicted by the model was at the low end of differences found empirically. The assumption that \( ps \) increases with prior study for low но high-frequency words produces a larger difference, large enough to allow the counter model to fit the data quite well.

4 Bowers (1999) suggested that with dissimilar alternatives in forced choice, participants rely on sublexical features to make a decision. Suppose that this is correct. Then, as flash time for the target is decreased, perception of these features will become progressively more difficult and eventually impossible. At that point, the decision process would have to switch to the same process as that used with similar alternatives: a word level process, according to Bowers, a process that produces a bias effect from prior study. Thus, at the fastest flash times (or zero flash time), bias should appear with dissimilar alternatives just as it does with similar alternatives. However, it does not, and so we conclude that Bowers’ hypothesis is wrong, that the decision process with dissimilar alternatives does not rely on sublexical features.
an already existing mechanism: the parameter $p_s$ specifies the quality of perceptual information from the stimulus. In earlier work, $p_s$ was affected by flash time for the stimulus: Longer flash times meant better information. Here, the quality of information for a low-frequency word increased with prior presentation. The reason for this result could be that the unfamiliarity of a low-frequency test alternative can disrupt processing on its first presentation or that perceptual information for a low-frequency word is better extracted on its second presentation. Either way, the new finding is accommodated without a change in the basic structure of the model, providing added confidence that the model captures some real aspects of performance.

Quantitative models provide a framework for understanding an experimental domain by bringing order and structure to experimental phenomena. Sometimes a model can be falsified by empirical tests that examine the fundamental structures of the model. For example, the memory models recently developed by Shiffrin and Steyvers (1997) and by McClelland and Chappell (1999) were designed in response to failures of earlier global memory models to pass such structural tests. More often, however, new data do not address a model’s fundamental structure. This was the case with Bowers’ (1999) efforts with respect to the counter model. Differences in instructions are easily explained in terms of the strategies that participants bring to a task. The finding that prior study of low-frequency words produces a genuine improvement in performance in forced choice contradicts claims made on the basis of earlier data (Ratcliff & McKoon, 1997), but the finding can be accommodated given a new understanding of how the perceptual system might have difficulty providing information about low-frequency words to the counters of the decision system. This does not change the integral basic structure of the model. In sum, the counter model exemplifies how a quantitative model can provide a framework that allows regularities in experimental data across tasks and conditions to be examined and hypotheses about the regularities to be tested.

References


Received November 22, 1998
Revision received July 26, 2000
Accepted August 21, 2000