Psychonomic Bulletin & Review
2000, 7 (4), 718-726

A criterion-shift model for enhanced discriminability in perceptual identification: A note on the counter model

ERIC-JAN M. WAGENMAKERS
and RENÉ ZEELENBERG
University of Amsterdam,
Amsterdam, The Netherlands

LAEL J. SCHOOLER
Pennsylvania State University,
University Park, Pennsylvania

and

JEROEN G. W. RAAIJMAKERS
University of Amsterdam,
Amsterdam, The Netherlands

The original version of the counter model for perceptual identification (Ratcliff & McKoon, 1997) assumed that word frequency and prior study act solely to bias the identification process (i.e., subjects have a tendency to prefer high-frequency and studied low-frequency words, irrespective of the presented word). In a recent study, using a two-alternative forced-choice paradigm, we showed an enhanced discriminability effect for high-frequency and studied low-frequency words (Wagenmakers, Zeelenberg, & Raaijmakers, 2000). These results have led to a fundamental modification of the counter model: Prior study and high frequency not only result in bias, but presumably also result in a higher rate of feature extraction (i.e., better perception). We demonstrate that a criterion-shift model, assuming limited perceptual information extracted from the flash as well as a reduced distance to an identification threshold for high-frequency and studied low-frequency words, can also account for enhanced discriminability.

When subjects identify briefly flashed words, their performance is affected if they have studied the flashed word previously, even though they have no explicit recollection of the study episode. This effect of long-term priming in implicit memory has attracted attention from several theorists (Bowers, 1999; Masson & Bodner, 2000; Masson & MacLeod, 1996; Ratcliff & McKoon, 1997; Schacter, 1994). Ratcliff and McKoon (1997) proposed an original and quantitative account of such implicit memory phenomena. Using a two-alternative forced-choice paradigm (e.g., LIED is briefly flashed and masked, followed by presentation of the response alternatives LIED and DIED), Ratcliff and McKoon claimed that subjects tended to prefer the studied alternative, regardless of whether that alternative had been flashed or not. Such a tendency would lead to benefits when the target (e.g., LIED) had been studied, but to costs when the foil (e.g., DIED) had been studied. Further, they found that the size of the benefits about equaled the size of the costs (e.g., Ratcliff, Allbritton, & McKoon, 1997; Rouder, Ratcliff, & McKoon, 2000). Hence, the effect of prior study was supposed to reflect a *bias* rather than some kind of *enhanced perceptual processing* of the flashed word.[1] Additional evidence for this idea came from the observation that even if performance was at chance when neither alternative was studied, effects of prior study were still present and thus appeared to be independent of information extracted from the flashed stimulus. Similarly, subjects had a preference to choose a high-frequency (HF) alternative such as MILE over a low-frequency (LF) alternative such as TILE. Therefore, effects of word frequency were likewise attributed solely to *bias*. Importantly, the biases for prior study and word frequency are supposedly mediated by different mechanisms, an issue debated by Wagenmakers, Zeelenberg, and Raaijmakers (2000). We will return to this later. For ease of reference, we will term the original version of the counter model the *counter model I*.

The counter model I is one of the few models to provide a quantitative account of repetition priming effects in visual word identification. The model successfully accounted for the effects of prior study and word frequency in three different word identification tasks: naming (or free response identification), forced-choice identification, and yes–no identification. Recent studies by Bowers (1999) and Wagenmakers et al. (2000), however, have shown that the counter model made some incorrect predictions. First, in a two-alternative forced choice paradigm, a choice between two HF alternatives was found to be more accurate than a choice between two LF alternatives. Second, prior study of *both* alternatives improved performance, albeit only for LF words. These results suggest problems for the counter model I, because it does not predict that prior study and word frequency affect the subject's ability to discriminate between the target and foil stimulus. In order to explain the enhanced discriminability effect for studied and HF words, McKoon and Ratcliff (in press; Ratcliff & McKoon, 2000) proposed a modification of the counter model. The new version of the counter model assumes that HF words as well as studied LF words, apart from having an advantage due to bias, also have a higher value of *ps*. The parameter *ps* denotes the probability of detecting information that enables one to discriminate between the target and the foil (e.g., the first letter of the LIED–DIED pair). We will term this modified model the *counter model II*.

Recently, two-alternative forced-choice procedure advocated by Ratcliff and McKoon (e.g., Huber, Shiffrin,

René Zeelenberg was supported by a grant from the Foundation for Behavioral and Social Sciences of the Netherlands Organization for Scientific Research. We thank Rich Shiffrin for helpful discussions and comments on an earlier draft of this paper. Correspondence concerning this article can be addressed to E.-J. M. Wagenmakers or R. Zeelenberg, Department of Psychonomics, Roetersstraat 15, 1018 WB Amsterdam, The Netherlands (e-mail: pn_wagenmakers@macmail.psy.uva.nl or pn_zeelenberg@macmail.psy.uva.nl).

Lyle, & Ruys, in press) has been adopted in several priming studies. One of the main advantages of this procedure is that it allows one to distinguish effects of enhanced discriminability from effects due to a simple bias to prefer the studied alternative, something difficult if not impossible to do in free response tasks. Claims that prior study causes "enhanced processing" are unwarranted when only free response measures are used. In fact, there is some evidence, albeit preliminary, that bias and enhanced processing have different decay functions (see McKoon & Ratcliff, in press) and may reflect different processes. The best way that we have found with which to assess the relative contributions of bias and enhanced processing involves use of forced-choice procedures.

In this paper, we will first consider the differences between the counter models I and II in more detail. We believe the counter model II represents a conceptual shift away from the original counter model. We will go on to argue that the different bias mechanisms that the counter model posits for prior study and word frequency are not mandated by the data currently available. Next, we will discuss the nature of enhanced discriminability as assumed by the counter model II—namely a higher rate of feature extraction for HF and studied LF words. We should like to point out that this is just one of the many ways to model enhanced processing. We will enumerate a number of established alternatives. Subsequently, we will present a new criterion-shift model and fits to the data from the Wagenmakers et al. (2000) study. The criterion-shift model was developed in response to the introduction of counter model II in order to demonstrate that the conceptual shift away from the original counter model might not be necessary. The criterion-shift model is very similar to the counter model II but does not need to assume an increase in the rate of feature extraction to model enhanced discriminability. Therefore, we believe that the criterion-shift model is conceptually more related to the counter model I than is the counter model II. After presenting the fits of the criterion-shift model to the data we will show that the assumptions of the criterion-shift model have ample precedent in the literature. The main point of the criterion-shift model is to illustrate that it is not *necessary* for prior study and/or word frequency to result in the extraction of more perceptual features. Instead, "decisional," bias-like processes may tune the word recognition process toward likely events, essentially mimicking the effects that a higher rate of feature extraction would have.

Differences Between the Counter Models I and II

In several articles, Ratcliff, McKoon, and co-workers have stressed that prior study merely leads to a bias in processing instead of some form of enhanced processing (Ratcliff et al., 1997; Ratcliff & McKoon, 1995, 1996, 1997; Ratcliff, McKoon, & Verwoerd, 1989). However, evidence for enhanced processing has now been consistently found (Bowers, 1999; Wagenmakers et al., 2000; Zeelenberg, Wagenmakers, & Raaijmakers, 2000). In order to handle the new data showing an overall increase in performance for studied words, the counter model II postulates two different mechanisms for prior study, one responsible for bias and one causing an increase in performance due to a higher rate of feature extraction. Regardless of the degree of change in the equations, this represents a considerable conceptual shift from the original counter model. Further, the counter model II differs from the counter model I in three assumptions. The new assumptions are as follows: (1) HF words are perceived better than LF words, (2) studied words are perceived better than nonstudied words, and (3) the second assumption holds only for LF words. This paper presents a criterion-shift model that does not need to assume a higher rate of feature extraction to model an overall increase in performance in 2-AFC perceptual identification.

Different Biases Due to Prior Study and Word Frequency

An important property of the counter model is that it has two different types of bias, one for prior study and one for word frequency. The bias due to prior study is best characterized as a *processing bias*: At each time step, the process is biased toward interpreting ambiguous information as evidence for the studied alternative. In contrast, the bias due to word frequency is a *resting-level* bias: HF words have a higher resting level of "counts" and thus have a head start in the comparison process. The rationale behind these two biases is the following: Word frequency effects are supposed to be unconditional or ubiquitous, whereas effects of prior study interact with similarity of the alternatives.[2] Prior study of one alternative leads to a bias for similar pairs (e.g., LIED–DIED), but not for dissimilar pairs (e.g., LIED–SOFA). The counter model posits that the processing bias is rather weak and only able to exert its effects for similar words that are presumably stored close together in memory. It was originally thought that if some aspect of the *memory representation* would be altered by prior study, such as the heightening of a resting level of activation, this would lead to the prediction that a bias effect should also be apparent for dissimilar pairs.[3]

In sum, the argument given above appears to preclude modeling prior study *in the counter model* as an increase in resting level of activation, since that would mean that a bias due to prior study would also manifest itself for dissimilar alternatives such as LIED–SOFA. But what is the argument against the bias due to word frequency's being a processing bias instead of a resting level bias? If word frequency bias is not a resting level bias but rather a processing bias, this would mean that *at each time step* the process is biased toward interpreting ambiguous information as evidence for the HF alternative. The bias due to word frequency would then be of the same nature as the bias due to prior study. It appears that the counter model is unable to address this issue when only accuracy is considered. To see why this is the case, consider the equation for probability correct in the counter models (this equation is discussed in the Appendix of Wagenmakers et al., 2000):

$$P(\text{correct}) = 1 - \{[(q/p)^a - (q/p)^z]/[(q/p)^a - 1]\},$$

where q/p is influenced by prior study and z is influenced by word frequency (for mathematical details see, e.g., Feller, 1968). Every change in P(correct) due to a change in z might just as well be accomplished by a change in q/p, since we have two variables that can be manipulated and only one equation. We know of no published data that urge separate bias mechanisms for word frequency and prior study.[4] For simplicity, our new model, which we will present shortly, therefore assumes that bias due to word frequency as well as bias due to prior study have the same locus. We will now turn to perhaps the most important feature of the counter model II, which is a higher rate of feature extraction for HF and studied LF words than for unstudied LF words.

On the Nature of Enhanced Discriminability: More Perceptual Features or More Efficient Processing?

Why are we better at identifying words that we have seen previously? The obvious answer would be to assume that certain *perceptual mechanisms* have "learned" from previous experience with the same stimulus. Such learning would allow a more rapid encoding as a result of prior study or high word frequency. Thus, the rate of feature extraction is supposedly higher for HF words or studied LF words, as the counter model II asserts. Under the same impression, we described the effects of enhanced discriminability as "*perceptual* gain" (Wagenmakers et al., 2000).

But the locus of the improvement does not necessarily have to be in perception per se. A word recently encountered might make its memory representation more readily available or more unitized in comparison with other memory representations that may be subject to some form of decay or interference. The advantages of unitization in perceptual tasks have been studied extensively. The finding that letters in words are more accurately identified than letters in nonwords (i.e., the word superiority effect) has been the inspiration for some well worked out models potentially capable of modeling improvement resulting from prior study and/or word frequency.

An in-depth discussion of the various ways of modeling enhanced discriminability resulting from unitized memory representations is clearly beyond the scope of this paper. Instead, we will briefly enumerate some possibilities. Enhanced discriminability might result from (1) top-down feedback from higher order units (e.g., McClelland & Rumelhart, 1981); (2) multiple independent sources of evidence—that is, low level perceptual information as well as information from higher order units (e.g., Massaro & Cohen, 1991; Norris, McQueen, & Cutler, 2000); (3) advantages of chunking (Anderson & Lebiere, 1998; Richman & Simon, 1989); (4) a selective bias toward automatic activation of familiar letter clusters (The better recognition of letters in words . . . is attributed to their advantage in competition for access to working memory, which is conferred by the encoding of familiar letter groups as units"; Estes & Brunn, 1987, p. 411); or (5) higher order units preventing encoded perceptual features from decaying or perturbing (Estes, 1975).

Obviously, then, the advantages of prior study and word frequency do not necessarily have to be due to some form of low-level perceptual learning. Enhanced discriminability might come about not through enhanced quality or quantity of extracted perceptual features, but rather through a more readily available higher order memory representation of the repeated stimulus. This means that highly available, unitized memory representations allow for more *efficient* processing of extracted perceptual information.

In the next sections, we will introduce a model that is in many aspects very similar to the counter model II. However, this model, the criterion-shift model, does not assume a change in the rate of feature extraction in order to handle our data showing an enhanced discriminability effect. Instead, it uses varying response criteria and the concept of limited information. First, we will present the new assumptions of the criterion-shift model in a general fashion. Next, we will show the fit of the model to the data from Wagenmakers et al. (2000). Finally, we will discuss the new assumptions underlying the criterion-shift model in more detail.

Explaining Enhanced Discriminability Without Assuming a Higher Rate of Feature Extraction: The Criterion-Shift Model

As mentioned before, the concept of *bias* was of central importance in the first counter model. In thinking about ways to adapt the counter model I to accommodate the enhanced discriminability effects for studied and HF words, but preserve some notion of bias, we eventually arrived at the following model. This model, the criterion-shift model, is very similar to the older counter models. However, the criterion-shift model demonstrates that enhanced discriminability can be modeled by a combination of threshold settings and the assumption of limited information extracted from the flash, instead of by an increase in the rate of feature extraction. We believe that this new model is conceptually more in line with the counter model I than is the counter model II. The new model also shows that the verbal labels "bias" and "enhanced processing" are not completely satisfactory to describe its processes. The criterion- shift model differs from the counter models I and II in the following three assumptions.

The first new assumption of the criterion-shift model is that information extracted from the flash is limited. The counter models I and II assume that the probability of extracting new diagnostic information from the briefly presented word does not diminish over time. For short periods of time after the flash, this is perhaps a good approximation, but for longer periods of time, it is certainly not. To incorporate the notion of limited information, we

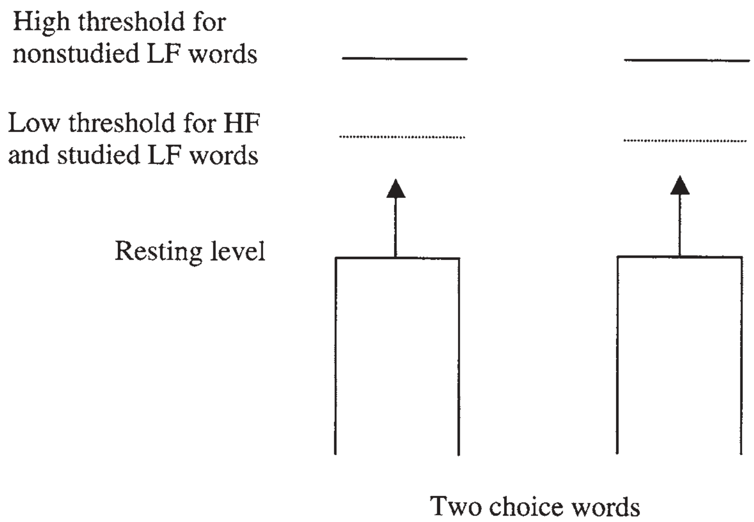

Figure 1. Illustration of the criterion-shift model as instantiated in a simple accumulator model. The two counters race toward their identification threshold that is located closer to the resting level for high-frequency (HF) and studied low-frequency (LF) word counters than for nonstudied LF word counters. If neither threshold is reached after *N* features have been evaluated, the subject guesses randomly between the two alternatives. Therefore, guessing is more likely for unstudied LF words, leading to a decrement in performance relative to that for HF and studied LF words.

imposed a deadline. If the process has not reached either identification threshold after all available units of information, say *N*, have been evaluated, the subject guesses randomly between the two alternatives (see also Schooler, Shiffrin, & Raaijmakers, in press).

The second new assumption of the criterion-shift model is that representations for HF and studied LF words require less evidence in order to reach their identification threshold than do representations for nonstudied LF words. This assumption is also present in other models, most predominantly in Morton's (1969) logogen model and the interactive activation model (McClelland & Rumelhart, 1981). Some models, such as the logogen model, propose that HF and studied LF word representations have a *lowered threshold* in comparison with that of nonstudied LF word representations. Other models, such as the interactive activation model, assume that the *resting level* of activation for HF and studied LF word representations is heightened in comparison with that of unstudied LF word representations. The difference between a higher resting level versus a lowered identification threshold is of no importance for the present application of the criterion-shift model. In fact, for the present model predictions, the two approaches (i.e., higher resting level vs. lowered threshold) are mathematically equivalent.

A third assumption concerns the decision rule. The criterion-shift model uses an absolute response criterion, whereas the counter models I and II both use a relative response criterion. This means that in the criterion-shift model, as in Morton's (1969) logogen model, an identification threshold is assumed to exist for each representation. At each point in time, one unit of information is attributed to one of the two representations or counters. The alternative whose representation crosses its own threshold first is chosen. Because the criterion-shift model uses an absolute response criterion, it is part of a larger class of models termed *simple accumulator* models. In contrast, the counter models I and II do not posit separate identification thresholds for each representation, but rather evaluate the amount of support for each alternative with respect to the amount of support for the competing alternative. As in the simple accumulator models, one unit of information is attributed to one of the two counters at each point in time. The response rule is that the counter first to be ahead by some criterion amount is chosen. Because the counter models I and II use a relative response criterion, they are part of a separate class of models—namely, *random walk* models. Metaphorically, the difference between the random walk models and the accumulator models can be illustrated by imagining a *race* between the representations of the two alternatives. In the random walk models, the representations continue to race until one representation is ahead by some criterion amount. In the accumulator models, the race ends once one of the representations crosses a finish line.

The Criterion-Shift Model and Its Fit to the Data

Just as the counter models I and II, the criterion-shift model assumes that words are represented by counters. Figure 1 shows the counters of the two response alternatives. Information will accumulate in the two counters until one of the two counters crosses its identification

Table 1
Observed Proportion of Correctly Identified Targets as a Function of Word Frequency and Study Status (Wagenmakers et al., 2000) and the Predictions of the Criterion-Shift Model

			Studied Alternative			
Target	Foil		Target	Foil	Neither	Both
HF	LF	Observed	.870	.773	.868	.842
		Predicted	.884	.786	.848	.834
LF	HF	Observed	.842	.683	.757	.793
		Predicted	.849	.685	.739	.804
HF	HF	Observed	.862	.753	.816	.822
		Predicted	.862	.770	.819	.819
LF	LF	Observed	.874	.720	.765	.821
		Predicted	.873	.702	.770	.819

Note—HF, high-frequency word; LF, low-frequency word.

threshold. The alternative whose counter crosses its identification threshold first is chosen (cf. the third assumption of the criterion-shift model mentioned in the previous section). However, only a limited amount of information is assumed to be extracted from the flash. Once all available information has been attributed to the counters but neither one has crossed its identification threshold, the subject will guess randomly between the two alternatives (cf. the first assumption of the criterion-shift model). As is apparent from Figure 1, the resting level of HF and studied LF words is closer to the identification threshold than is the resting level of unstudied LF words (cf. the second assumption of the criterion-shift model). Further details of the model are outlined in the Appendix.

The data of Wagenmakers et al. (2000) consisted of 16 conditions. Four levels of prior study (i.e., the target, the foil, neither alternative, or both alternatives) were combined with four levels of word frequency of the alternatives (i.e., HF–LF, LF–HF, HF–HF, and LF–LF, where the first member of the pair denotes the target and the second denotes the foil). Only similar word pairs (e.g., LIED–DIED) were used. The data of Wagenmakers et al. as well as the predictions of the criterion-shift model are shown in Table 1. As can be seen, the fit to the data is excellent. For 9 of the 16 conditions, the model was off by less than 1% and for no condition was the model off by more than 2%. We should especially like to point out that the model does a good job of predicting improvement in performance resulting from prior study of both LF alternatives. The model would *not* make this prediction if information extracted from the flash was unlimited (i.e., $N \to \infty$, which would mean that there would be an endless amount of information to be evaluated and no guessing would occur). This is because identification thresholds located close to the resting level of activation are more likely to be arrived at by chance fluctuations, leading to a *decrement* in performance. However, in the criterion-shift model, the costs of more often crossing the identification threshold of the incorrect alternative are strongly counteracted by the benefits of having to guess less often.

On the New Assumptions of the Criterion-Shift Model

In comparison with the counter model I, the criterion-shift model uses three new assumptions in order to fit the data of Wagenmakers et al. (2000). The first assumption is that information is limited and subjects will *randomly* guess between the two alternatives once all available information is depleted without having led to a decision (i.e., by not arriving at either identification threshold on the basis of the extracted perceptual information). A similar proposal has been made by Anderson and Lebiere (1998) with respect to an ACT-R model for the word superiority effect:

> subjects perceive an object and do not have access to the basis of their perception. Thus, given the top half of the D, a subject might perceive R. Thinking they perceived R and given a choice between D and K they have no option but to guess; they do not have access to the features that gave rise to the R perception and that would allow them to choose between D and K. (p. 161)

Another *pure-guess model* was proposed by De Jong (1991) to account for signal-to-respond data, cautioning against the use of the speed–accuracy decomposition technique that supports a sophisticated-guessing model (Meyer, Irwin, Osman, & Kounios, 1988): "According to the pure-guess model, guessing processes do not have access to preliminary, partial output of regular processes; consequently, guesses will be at chance accuracy" (De Jong, 1991, p. 335).

In general, pure-guess models are less optimal than sophisticated-guessing models, since the former type of model *discards* potentially useful information if the amount of information does not exceed a prerequisite quantity needed to pass threshold. The property of suboptimality, inherent in pure-guess models, is related to the notion of a threshold. A threshold can be thought of as a demarcation of a discrete transition from one state (e.g., unidentified) to another (e.g., identified). The suboptimality property is in fact crucial for the explanation of enhanced discriminability without the assumption of a higher rate of feature extraction. A sophisticated-guessing model would fail to predict the improvement in performance due to prior study of both LF alternatives. In a sophisticated-guessing model, the system would simply choose the alternative with the nearest response boundary, thereby using all available information and making an optimal decision. In this situation, there would be no benefit of having to guess less often for representations with response boundaries located close to the starting point, since the guesses are *informed guesses* and will not be at chance accuracy. Rather, as for the case when unlimited information is assumed (i.e., $N \to \infty$, mentioned above), response boundaries located close to the starting point would only hurt performance because of a higher vulnerability to chance fluctuations. Thus, it turns out that the suboptimality property inherent in the pure-guess assumption of the criterion-shift model is vital for the prediction of enhanced discriminability.

One might wonder why any system, and especially a well-trained system such as that involved in word recognition, would develop to be suboptimal rather than optimal. In other words, is the kind of suboptimality proposed here merely a limitation of the system, or is there any adaptive value in *not* having access to all features that give rise to the perception of higher order units? It would seem to us that the ignorance of a system with respect to its basic perceptual processes (i.e., a pure-guess model) might very well be functional. The computational resources needed for a system that only considers above-threshold activated concepts would be much less than those needed for a system that has to consider the contribution of every low-level perceptual feature to the activation of the corresponding representations. Disregarding *potentially* valuable but noisy information might have useful filter-like properties in a complex world.

The second assumption that distinguishes the criterion-shift model from the counter model II is that HF and *studied* LF words require less evidence to be identified than do unstudied LF words. Morton (1969; see also Broadbent, 1967; McClelland & Rumelhart, 1981) advanced the same idea, proposing that representations for HF and studied LF words need less information in order to pass an identification threshold than do representations of unstudied LF words. In such models, it is generally assumed that every time a word is encountered, the amount of evidence needed for its representation to pass threshold on a subsequent encounter is decreased according to some nonlinear function (e.g., a negatively accelerated function). In the present work, we make the simplifying assumption that in contrast to LF words, HF words are already at asymptote and prior study will therefore result in a negligible decrease of the distance to the threshold for HF words.[5] The reduction of the distance to the threshold as a result of prior exposure can be modeled either through a heightening of the level of activation or through the lowering of the response threshold for the repeated word representation. As mentioned in one of the preceding sections, for the present application of the criterion-shift model both methods are mathematically equivalent.

The third assumption that distinguishes the criterion-shift model from the counter models I and II is the use of an absolute rather than a relative response criterion. Recently, Ratcliff and McKoon (1997, p. 336) pointed out that models using an absolute response criterion such as the logogen model (Morton, 1969) incorrectly predict that a forced choice between HF words is less accurate than a forced choice between LF words, because of a higher impact of chance fluctuations for HF words. Since HF words need less evidence to cross their identification threshold, they are more likely to do so in error. This analysis is correct when one assumes that the information extracted from the flash is unlimited and continuous over time (i.e., $N \to \infty$). However, when one assumes that the information extracted from the flash is limited, and subjects guess when all extracted information is used before a response criterion is satisfied, the need for less evidence in order to cross an identification threshold might be beneficial. When less information is needed in order to cross the threshold, this reduces the proportion of guesses, improving performance.

As a final note, we would like to mention the following. One could claim that the enhanced discriminability effect is really a kind of bias in the models presented here, since it comes about through adjustment of the identification threshold. An adjustment of identification thresholds is typically considered a form of bias (e.g., Broadbent, 1967). However, one could also argue that HF and studied LF words are processed more efficiently than nonstudied LF words. Because subjects have to guess less often for HF and studied LF words, their responses are more often based on the information extracted from the briefly presented stimulus. The latter interpretation would be more consistent with an enhanced processing view than with a bias view. In any case, enhanced discriminability does not arise in the criterion-shift model, because more perceptual information is extracted from the stimulus per unit time.

Discussion

Our aim in the present paper has been to show that enhanced discriminability does not necessarily *imply* a higher rate of feature extraction. This is not a trivial point. In contrast to free response procedures, 2-AFC procedures are often used to control for "response bias" or "threshold bias." When one finds an improvement in accuracy (i.e., enhanced discriminability) while these controls are in place, it is tempting to conclude that the flashed stimulus is encoded faster. The general implication of the model presented here is that the locus of enhanced discriminability in forced choice does not *necessarily* lie at an early perceptual stage of processing, even though a simple response bias or preference can be controlled for.

The demonstration that enhanced discriminability can be accounted for without an enhanced rate of feature extraction does not show that the counter model II is wrong. It might very well be that the rate of feature extraction for repeated stimuli is higher than that of newly encountered stimuli. The criterion-shift model presented here gives an *alternative* explanation for the enhanced discriminability effect. Therefore, it would be of considerable theoretical interest to be able to distinguish the counter model II, positing a higher rate of feature extraction for HF and studied LF words, from the criterion-shift model. One possible way to distinguish the two types of models would be to use *confidence ratings*. In the criterion-shift model, the subject can be in either one of the following two states: (1) the threshold-passing stage that will presumably lead to highly confident responding, or (2) the guessing state in which the subject is not confident and performance is at chance. If the confidence ratings are bimodally distributed and the accuracy of the component reflecting unconfident responses is at chance, this would constitute very strong support for the criterion-shift model and pose severe problems for the counter model II.

We believe that the criterion-shift models presented here could also explain the result recently obtained in short-term priming—namely, that a prime semantically related to both alternatives enhances performance in the subsequent 2-AFC test (Huber et al., in press). To accommodate this finding, one would have to assume that the presentation of a semantically related prime (e.g., BANANA) reduces the distance to the threshold of the two alternatives (e.g., KIWI–PEAR). Such an assumption was made by Morton (1969), for example. In addition, we should like to draw attention to a related debate on the impact of context on perceptual encoding. Several studies (e.g., Masson & Borowsky, 1998; Reinitz, Wright, & Loftus, 1989; Rhodes, Parkin & Tremewan, 1993) have shown that a semantically related prime can enhance a signal detection measure of performance such as d' across a range of different tasks. This has been taken to suggest that the semantically related prime increases the rate of feature extraction (Reinitz et al., 1989), and that the d' effect reflects a change in perceptual processing (Farah, 1989). If this were unambiguously true, it would make the models presented here less plausible, since the semantic overlap is 100% for repeated words. However, the criterion bias model presented by Norris (1995) as well as the connectionist model presented by Masson and Borowsky both show that the d' effect can in fact be modeled *without* assuming perception to be directly affected by the semantically related prime. These two models have not been applied to the enhanced discriminability effects in a 2-AFC paradigm where a simple response bias can be controlled.

Finally, we would like to discuss how the criterion-shift model fares against two important findings in the literature—that is, (1) the bias effect of word frequency in the absence of a to-be-identified stimulus, and (2) the absence of a bias effect due to prior study when the choice alternatives are dissimilar (e.g., LIED–SOFA). First, in free response identification, it has been found that even in the absence of a stimulus, subjects tend to "identify" HF words rather than LF words (for a discussion, see Broadbent, 1967). Both the counter model II and the criterion-shift model can account for this phenomenon. In the counter model II, word frequency has two effects: a resting level bias and a higher rate of feature extraction for HF words. When no stimulus is flashed, there is no valid perceptual information for discriminating the choice alternatives (i.e., parameter *ps* equals zero). The resting level bias will, however, still be operative and lead to a preference for HF words over LF words. In the criterion-shift models, word frequency also has two effects: a processing bias and a threshold bias (i.e., HF words need less information in order to cross an identification threshold). In the absence of a stimulus, both biases will favor "identification" of the HF alternative. Second, in this paper, we have focused on similar pairs of alternatives (e.g., LIED–DIED). The empirical state of affairs for *dissimilar* alternatives is still very much in flux. As noted before, Ratcliff and McKoon (1997) found that prior study had no effect for dissimilar alternatives. However, Bowers (1999) found bias effects due to prior study both for pairs of similar and for pairs of dissimilar alternatives. In response, McKoon and Ratcliff (in press) argued that Bowers's instructions might have led subjects to use explicit-retrieval strategies. Explicit-retrieval strategies are expected to cause bias for both similar and dissimilar pairs of alternatives. To complicate matters even further, Masson (2000) recently obtained bias effects due to prior study exclusively for pairs of similar alternatives, although his instruction resembled that of Bowers. In the counter model II, prior study has two effects: (1) a processing bias, operative only for similar pairs of alternatives, and (2) a higher rate of feature extraction for studied LF target words. Therefore, the counter model II predicts that prior study results in both bias and enhanced discriminability for *similar* LF alternatives, but only in enhanced discriminability for *dissimilar* LF alternatives. In the criterion-shift model, the assumption about the higher rate of feature extraction for a studied LF word is replaced by the adjustment of the identification threshold. Note that the counter model II assumes that the higher rate of feature extraction for studied LF words comes into play only when the studied LF word is also the target alternative. In contrast to the counter model II, the criterion-shift model adjusts the threshold for any studied LF alternative, regardless whether that alternative is the target or the foil. Thus, the criterion-shift model *theoretically* predicts that for dissimilar LF alternatives, prior study of the foil alternative leads to a decrement in performance. However, it turns out that when the parameter *ps* (i.e., the probability of observing discriminative information) is as high as in the simulations reported here, a small adjustment of the threshold for the foil alternative is inconsequential with respect to the behavior of the model. In other words, although the theoretical mechanism underlying prior study of one dissimilar LF alternative leads to both costs after study of the foil and benefits after study of the target, the predicted data will show an improvement only as a result of studying the target and a negligible decrement in performance due to study of the foil. Therefore, both the counter model II and the criterion-shift model will make quantitatively similar predictions concerning the interaction between prior study and the similarity of the alternatives.

Summary and Conclusions

The counter model I (Ratcliff & McKoon, 1997) was a highly original account for implicit memory phenomena. Moreover, it was a quantitative and testable model, and it provided a successful account of a wide range of results available at that time. A series of recent findings (Bowers, 1999; Wagenmakers et al., 2000; Zeelenberg et al., 2000), however, have shown that contrary to the predictions of the original counter model, prior study not only results in bias, but also results in enhanced discriminability. In the present paper, we have shown that there are many ways to explain enhanced discriminability, and we have presented a simple accumulator model to fit the data from Wagenmakers et al. This new criterion-shift model does not need to assume a higher rate of feature extraction or some other "perceptual" gain to account for enhanced discriminability effects for HF and studied LF words. Rather, it turns out that by varying response criteria and by limiting the information extracted from the flash, one can obtain these effects also.

REFERENCES

Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Erlbaum.

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

Broadbent, D. E. (1967). Word-frequency effect and response bias. *Psychological Review*, **74**, 1-15.

De Jong, R. (1991). Partial information or facilitation? Different interpretations of results from speed–accuracy decomposition. *Perception & Psychophysics*, **50**, 333-350.

Estes, W. K. (1975). The locus of inferential and perceptual processes in letter identification. *Journal of Experimental Psychology: General*, **104**, 122-145.

Estes, W. K., & Brunn, J. L. (1987). Discriminability and bias in the word-superiority effect. *Perception & Psychophysics*, **42**, 411-422.

Farah, M. (1989). Semantic and perceptual priming: How similar are the underlying mechanisms? *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 188-194.

Feller, W. (1968). *An introduction to probability theory and its applications*. New York: Wiley.

Huber, D. E., Shiffrin, R. M., Lyle, K. B., & Ruys, K. I. (in press). Perception and preference in short-term word priming. *Psychological Review*.

Luce, R. D. (1986). *Response times*. New York: Oxford University Press.

Massaro, D. W., & Cohen, M. M. (1991). Integration versus interactive activation: The joint influence of stimulus and context in perception. *Cognitive Psychology*, **23**, 558-614.

Masson, M. E. J. (2000). *Bias effects in repetition priming of masked word identification: Differential influences of modality and generation*. Manuscript submitted for publication.

Masson, M. E. J., & Bodner, G. E. (2000, January). *Fluent encoding of probes guides masked word identification*. Paper presented at the 25th Annual Interdisciplinary Conference, Teton Village, WY.

Masson, M. E. J., & Borowsky, R. (1998). More than meets the eye: Context effects in word identification. *Memory & Cognition*, **26**, 1245-1269.

Masson, M. E. J., & MacLeod, C. M. (1996). Contributions of processing fluency to repetition effects in masked word identification. *Canadian Journal of Experimental Psychology*, **50**, 9-21.

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

McKoon, G., & Ratcliff, R. (in press). Bias in the counter model. *Psychological Review*.

Meyer, D. E., Irwin, D. E., Osman, A. M., & Kounios, J. (1988). The dynamics of cognition and action: Mental processes inferred from speed–accuracy decomposition. *Psychological Review*, **95**, 183-237.

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

Norris, D. (1995). Signal detection theory and modularity: On being sensitive to the power of bias models of semantic priming. *Journal of Experimental Psychology: Human Perception & Performance*, **21**, 935-939.

Norris, D., McQueen, J. M., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral & Brain Sciences*, **23**, 299-370.

Raaijmakers, J. G. W., Zeelenberg, R., Schooler, L. J., & Wagenmakers, E. M. (2000). *Repetition priming and enhanced discriminability in perceptual word identification*. Manuscript in preparation.

Ratcliff, R., Allbritton, D. W., & McKoon, G. (1997). Bias in auditory priming. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **23**, 143-152.

Ratcliff, R., & McKoon, G. (1995). Bias in the priming of object decisions. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 754-767.

Ratcliff, R., & McKoon, G. (1996). Bias effects in implicit memory tasks. *Journal of Experimental Psychology: General*, **125**, 403-421.

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

Ratcliff, R., & McKoon, G. (2000). Modeling the effects of repetition and word frequency in perceptual identification. *Psychonomic Bulletin & Review*, **7**, 713-717.

Ratcliff, R., McKoon, G., & Verwoerd, M. (1989). A bias interpretation of facilitation in perceptual identification. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 378-387.

Reinitz, M. T., Wright, E., & Loftus, G. R. (1989). Effects of semantic priming on visual encoding of pictures. *Journal of Experimental Psychology: General*, **118**, 280-297.

Rhodes, G., Parkin, A. J., & Tremewan, T. (1993). Semantic priming and sensitivity in lexical decision. *Journal of Experimental Psychology: Human Perception & Performance*, **19**, 154-165.

Richman, H. B., & Simon, H. A. (1989). Context effects in letter perception: Comparison of two theories. *Psychological Review*, **96**, 417-432.

Rouder, J. N., Ratcliff, R., & McKoon, G. (2000). A neural network model of implicit memory for object recognition. *Psychological Science*, **11**, 13-19.

Schacter, D. L. (1994). Priming and multiple memory systems: Perceptual mechanisms of implicit memory. In D. L. Schacter & E. Tulving (Eds.), *Memory systems 1994* (pp. 233-238). Cambridge, MA: MIT Press.

Schooler, L., Shiffrin, R. M., & Raaijmakers, J. G. W. (in press). Theoretical note: A Bayesian model for implicit effects in perceptual identification. *Psychological Review*.

Wagenmakers, E.-J. 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**, 662-667.

Zeelenberg, R., Wagenmakers, E. M., & Raaijmakers, J. G. W. (2000). *Repetition priming in implicit memory tasks: Prior study causes enhanced processing, not only bias*. Manuscript submitted for publication.

NOTES

1. We prefer to use *enhanced processing* to avoid confusion with the signal-detection term *sensitivity*, which refers to a unidimensional decision process.

2. Experiments testing the prediction of the counter models I and II that the preference for HF words over LF words is unaffected by similarity of alternatives have, to the best of our knowledge, not been reported. We are currently running experiments to test this prediction.

3. However, Schooler, Shiffrin, and Raaijmakers (in press) showed that this reasoning was incorrect, demonstrating that additional variability associated with a choice between dissimilar alternatives could

produce smaller bias effects than those observed for a choice between similar alternatives. In the following, we will ignore this issue and give the line of argumentation from a counter model framework.

4. When one does not focus solely on accuracy, but also takes response latency into account, this introduces a *second* equation of the counter model. Thus, by considering accuracy and response latencies simultaneously, one might be able to eliminate the identifiability problem with respect to the different biases posited for word frequency and prior study.

5. Note that Raaijmakers, Zeelenberg, Schooler, and Wagenmakers (2000) did obtain a small but robust enhanced discriminability effect for HF words, but only after these HF words were studied multiple times. This result is consistent with the assumption that the interaction of prior study with word frequency is of a continuous rather than a discrete nature.

APPENDIX
Mathematics of the Criterion-Shift Model

In this Appendix, we will present some details of the criterion-shift model for perceptual forced-choice, as instantiated in a simple accumulator model. Parameter optimization was conducted using the analytic expressions given here, and simulations confirmed the results.

Mathematically, the criterion-shift model is a simple accumulator model (see, e.g., Luce, 1986). Figure 1 provides an illustration of the simple accumulator model. Each unit of time, one unit of information (i.e., a count) is added to one of the two representations corresponding to the two response alternatives (i.e., counters). This accumulator model is called "simple" because it is discrete both in time and in the way information is accrued. The counter that crosses its identification threshold first is chosen. If neither the target counter nor the foil counter reaches its identification threshold by the time N units of evidence have been evaluated, the subject guesses randomly between the two alternatives. The first equation gives the probability p that a count is added to the counter of the target alternative:

$$p = ps + (1 - ps)\,(0.5 + \beta s + \beta f), \tag{1}$$

where ps is the probability of detecting target diagnostic information (e.g., the first letter of the LIED-DIED pair), βs is a bias to interpret ambiguous information (i.e., $1 - ps$) in favor of the studied alternative (βs is a positive number when only the target has been studied and a negative number when only the foil has been studied), and βf is a bias to interpret ambiguous information in favor of the HF alternative when the foil is LF (βf is a negative number when the target is LF and the foil is HF). Note that this frequency bias is no longer constrained to be a resting level bias, as in the counter models I and II, but instead also affects p (i.e., a processing bias). As we have already pointed out, for the counter models I and II this difference is inconsequential as long as only accuracy is considered. For the fits to the data from Wagenmakers et al. (2000) reported in Table 1, $ps = .1752$, $\beta s = .0199$, $\beta f = .0063$.

The criterion-shift model assumes that LF alternatives need slightly more evidence in their favor before reaching their identification threshold than do HF alternatives. In this case, we took 1 as the additional unit of information needed in order to arrive at the identification threshold for an LF word as opposed to an HF word. Prior study leads to LF words' needing less evidence to cross threshold, and we made the simplifying assumption that studied LF words need the same number of counts to cross threshold as HF words. Specifically, the distance from unstudied LF words to their identification threshold was set to 20. For HF and studied LF words, this distance was reduced by 1 to 19.

P_f is the probability of the incorrect (foil) counter reaching its identification threshold before the correct (target) counter. Let k_t denote the distance from the resting level of the target counter to its identification threshold, and let k_f denote the distance from the resting level of the foil counter to its identification threshold. Then

$$P_f = \sum_{j=0}^{k_t - 1} p^j q^{k_f} \binom{k_f + j - 1}{j} \tag{2}$$

is the formula used to calculate P_f without taking N (i.e., the cut-off due to guessing) into account. P_t, the probability of the target counter's reaching its identification threshold before the foil counter, can be obtained by replacing k_t with k_f and p with q in Equation 2. To arrive at the probability of the foil counter's reaching its threshold before the target counter as well as before N, where N equals the number of steps until guessing, the summation should only be applied when $j + k_f < N$. In the fits reported in Table 1, N was fixed at 35.

Finally, the formula for probability correct is the sum of the following two probabilities: (1) the probability that the target counter crosses its identification threshold both before N evaluations and before the foil counter crosses its identification threshold, and (2) the probability of correctly *guessing* the target alternative once neither the target nor the foil counter has reached threshold after evaluation of N units of information:

$$P(\text{correct}) = P_t + 0.5\,[1 - (P_t + P_f)]. \tag{3}$$

(Manuscript received March 16, 2000;
revision accepted for publication August 29, 2000.)