Running head: Models for priming

REMI and ROUSE: Quantitative Models for Long-Term and Short-Term Priming in Perceptual Identification

Eric-Jan M. Wagenmakers[1], René Zeelenberg[1], Dave Huber[2], Jeroen G.W. Raaijmakers[1], Richard M. Shiffrin[3], and Lael J. Schooler[4]

1 University of Amsterdam
2 University of Colorado
3 Indiana University
4 The Pennsylvania State University

The REM model originally developed for recognition memory (Shiffrin & Steyvers, 1997) has recently been extended to implicit memory phenomena observed during threshold identification of words. We discuss two REM models based on Bayesian principles: a model for long-term priming (REMI; Schooler, Shiffrin, & Raaijmakers, 1999), and a model for short-term priming (ROUSE; Huber, Shiffrin, Lyle, & Ruys, in press). Although the identification tasks are the same, the basis for priming differs in the two models. In both paradigms we ask whether prior study merely reflects a *bias* to interpret ambiguous information in a certain manner, or instead leads to *more efficient encoding*. The observation of a 'both-primed benefit' in two-alternative forced-choice paradigms appears to show that *both* processes are present. However, the REMI model illustrates that the both-primed benefit is not necessarily indicative of an increase in perceptual sensitivity but might be generated by a criterion bias. The ROUSE model demonstrates how the amount of attention paid to the prime, and the consequent effect upon decision making, may lead to the *reversal* of the normal short-term priming effect that is observed in certain conditions.

A stimulus in a current task is said to be *primed* when it has been encountered previously, but the memory for that prior occurrence is not required for performance of the current task. Primed stimuli are generally responded to faster and more accurately than unprimed stimuli. For instance, the earlier presentation of a word enhances probability of correctly identifying that word in a later identification task (Church & Schacter, 1994). It is important to note that this priming effect can come about without subjects being aware of any relation between the task used at test and the earlier study episode. Moreover, repetition priming effects can even be observed in amnesic patients who are unable to remember the study episode (Shimamura & Squire, 1984). Such an effect upon performance is termed an *implicit* memory effect. A large variety of tasks have been used to study implicit memory processes. Among these tasks are auditory and visual word identification (Bowers, 1999; Jacoby & Dallas, 1981; Masson & Freedman, 1990; Salasoo, Shiffrin, & Feustel, 1985), picture identification, word stem completion (Graf, Squire, & Mandler, 1984), word fragment completion (Roediger, Weldon, Stadler, & Riegler, 1992) and free association (Zeelenberg, Shiffrin, & Raaijmakers, 1999).

In this chapter, we will present two models of priming. The primary task under consideration is the identification of words presented visually at threshold. The first model, REMI (for Retrieving Effectively from Memory, the 'I' stands for implicit) developed by Schooler et al. (1999), is a model for *long-term priming* in implicit memory. It explains repetition priming effects by assuming that during study of a word some contextual information is added to the corresponding lexical trace. This contextual information stored during the study task will tend to match the contextual information present during the test task, leading subjects to prefer studied words over nonstudied words. The second model, ROUSE (Huber et al., in press), is a model of *short-term* priming. ROUSE stands for Responding Optimally with Unknown Sources of Evidence, and it is able to explain an intricate pattern of results. ROUSE assumes that (a) information arriving due to presentation of the primes can be mistaken for valid target information (i.e., *source confusion*), and (b) subjects *discount* evidence from such features to a degree determined by their estimate of this source confusion. ROUSE accounts for the finding that short-term repetition priming results in a bias to choose *repeated* words following the *passive* viewing of primes whereas the *active* processing of primes results in a bias to choose *non-repeated* words. Both REMI and ROUSE can be thought of as natural extensions of the REM model (for details of the REM model applied to episodic recognition memory see Shiffrin & Steyvers, 1997). The REM model is a general memory model based upon the idea that subjects make decisions in an optimal (i.e., Bayesian) fashion on the basis of noisy information. Although derived on a quite different basis, REM shares many similarities with the well-known SAM model for episodic memory tasks such as recognition and recall (e.g., Gillund & Shiffrin, 1984; Raaijmakers & Shiffrin, 1981). REMI and ROUSE demonstrate how priming effects in implicit memory can be explained in the framework of traditional information processing models (cf. Ratcliff & McKoon, 1997).

The priming models we will discuss here are *quantitative* models. They are based on mathematical expressions and provide detailed fits to a wide variety of empirical results. Until recently, most theory development in implicit memory has been *qualitative* in nature, that is, based on verbal descriptions. Although verbal labels are readily accessible to

researchers, they are by their very nature imprecise and notoriously susceptible to alternative interpretation.

The outline of this chapter is as follows. We first describe briefly the two-alternative forced-choice perceptual identification task that is used to explore short- and long-term priming, and sketch in broad terms the way we model performance in such tasks. We then discuss the nature of long-term priming in implicit memory tasks, present data from our basic task, and describe the REMI model and its fit to the data. We end by discussing short-term priming, data from our basic task, and finally the ROUSE model and its fit to the findings.

Two-Alternative Forced-Choice Perceptual Identification

The task used to explore both short- and long-term priming is quite simple: A word (termed a target) is presented visually for a brief period of time, and then followed by a mask, conditions that make perception difficult. This display is followed almost immediately by two choice words: the flashed *target* word and an incorrect *foil* word. The subject attempts to decide which choice word had been flashed. The use of a two-alternative procedure allows the set of alternatives and their similarity to be controlled, and allows the relation to prior primes to be precisely manipulated.

The REM-inspired modeling for such a task assumes that the result of perception of the initial display is a noisy vector of features. Each alternative is represented as a vector of features, albeit without noise. A comparison of the perceived vector to the choice vectors is carried out and an optimal, Bayesian-based, decision is made (e.g., in some, but not all, cases, the better matching alternative is chosen). Long- and short-term priming are assumed to have an effect by altering the degree of matching, as we describe in the sections below. One way to affect matching is to improve the accuracy of the perceived vector, which could be thought of as 'more efficient encoding'. A second way to affect matching is by incorporating features from the prime in the perceived vector, independent of the actual flash. This may be though of as 'bias'. The REMI and ROUSE models account for the data in terms of bias, proposing different bias-mechanisms for long- and short-term priming.

Long-Term Priming in Implicit Memory Tasks: Perceptual Improvement or Bias?

In long-term priming, the prime occurs a considerable time before the primary task. Generally, effects of long-term priming are only observed when the target is a repetition of a studied item (but see Becker, Moscovitch, Behrmann, & Joordens, 1997). The effect of long-term repetition priming in implicit memory tasks is usually interpreted as *perceptual facilitation* due to prior exposure (e.g., Salasoo et al., 1985; Squire, 1992). Schacter (1990) has argued that "visual priming may make it easier ... to extract visual information from the test cue" (p. 237). The general notion is that the perceptual analysis of a primed stimulus in the test task involves the same encoding processes employed when the stimulus was first seen in the study task. The encoding processes supposedly learn from the earlier encounter with the stimulus, resulting in a faster rate of feature extraction the second time the stimulus is presented. In this account, the priming effect is due to low level perceptual learning. This perceptual learning is supposedly dissociated from conscious awareness and its existence has been taken to suggest the presence of several separate memory systems, such as for example the visual word-form system (e.g., Schacter, 1990). Generally, however, the mechanisms underlying enhanced feature extraction are not clearly specified. Thus it remains somewhat unclear *why* or *how* prior study results in better performance. This makes it hard to derive testable predictions.

Ratcliff and McKoon (1997; Ratcliff, McKoon, & Verwoerd, 1989) have proposed a different account of repetition priming. According to Ratcliff and McKoon, prior study does not result in a more accurate perception of the repeated stimulus, but rather biases subjects to identify the stimulus as one they studied before. The most convincing evidence supporting their claim came from several experiments using the *2-AFC paradigm* for perceptual identification of visually presented stimuli. In a typical 2-AFC experiment, a word, say LIED, is briefly flashed and subsequently masked. Next, two alternatives appear: the target word LIED and an orthographically similar foil word, say DIED. Subjects have to choose which of the two alternatives was flashed. Suppose the target word LIED has been studied some time prior to the 2-AFC test (the typical study instruction is 'study these words for a later memory test'). Consistent with the traditional 'encoding advantage' view of priming, prior study of the target word *enhances* performance compared to the condition in which neither alternative was studied. Several studies, however, have shown that when the orthographically similar foil DIED has been studied, performance *decreases* (Ratcliff & McKoon, 1996, 1997; Ratcliff, Allbritton, & McKoon, 1997). Moreover, the costs in performance of studying the foil alternative about *equals* the gain or benefit due to studying the target alternative, leading Ratcliff & McKoon to characterize priming as *bias*. Two additional findings further supported the notion of bias. First, when flash time is so short that performance is at chance, or even when instead of a word some noise characters are flashed, the priming pattern of costs and benefits is left largely unaffected. Second, when the two alternatives are not orthographically similar to each other (e.g., LIED-SOFA) Ratcliff and McKoon (1997) found little or no effect of prior study, contradicting the claim that prior study results in enhanced processing (alternative findings by Bowers, 1999, will be discussed later).

Ratcliff and McKoon (1997) presented an elegant model to account for these data, the so-called *counter model*. The counter model assumes that perceptual information extracted from the flash can be of two kinds. Information that is helpful in making a correct decision (i.e., the first letter of a choice between LIED and DIED) is termed diagnostic and will always provide evidence in favor of the target alternative. Information that is nondiagnostic (whether it be the last three letters of the LIED-DIED pair or perceptual noise), will sometimes be taken to support the target alternative, and sometimes be taken to support the foil alternative. When neither alternative has been studied previously, nondiagnostic information is equally likely to support either the target alternative or the foil alternative. If one alternative has been studied however, this results in the studied alternative having a greater probability of acquiring the support of the nondiagnostic information than the nonstudied alternative. Thus, the counter model assumes ambiguous information is processed to be consistent with previous experience. Ratcliff and McKoon term this 'stealing' of information (i.e., the studied alternative steals nondiagnostic information from the nonstudied alternative). When it is additionally assumed that such stealing of information is assumed to occur only between similar alternatives (e.g., LIED-DIED) the model predicts the observed pattern of costs and benefits resulting from prior study. Since stealing also applies to perceptual noise, the counter model correctly predicts that the bias effect due to prior study should be present even when noise characters are flashed instead of a regular target word. In sum, the counter model explains repetition priming effects not through a higher rate of feature extraction, but through a biased allocation of available features.

As illustrated by the work of Ratcliff and McKoon outlined above, the 2-AFC paradigm can be used to distinguish between a target selective improvement in performance on the one hand and a simple processing bias on the other. Now consider what the different hypotheses of priming (i.e., the 'encoding-advantage' versus 'processing bias' view) predict for performance in the 2-AFC task when both alternatives are studied. The encoding-advantage view of priming predicts a target selective advantage such that performance will be better whenever the target has been studied. As compared to the condition where neither alternative was studied this will result in performance improvements when *both* alternatives were studied as well as when only the target was studied. In contrast, the counter model predicts no 'both-primed benefit'[1], since it assumes that the biases for the target and foil alternative cancel. Recent work has specifically addressed this issue and has consistently found a both-primed benefit with low-frequency words (as well as the usual bias effects). Figure 1 shows typical data from an auditory word identification experiment (Zeelenberg, Wagenmakers, & Raaijmakers, 2000, Experiment 2). The size of the both-primed benefit is reflected by the difference between the middle two bars in Figure 1. Although the bias effect appears to be larger than the both-primed benefit, the latter has been consistently found across a range of tasks such as visual word recognition (Bowers, 1999; McKoon & Ratcliff, in press; Wagenmakers, Zeelenberg, & Raaijmakers, in press), auditory word recognition, and fragment completion and picture identification (Zeelenberg et al., 2000). These data illustrate that there is more to priming than just a simple bias to prefer the studied word.

It is tempting to view the both-primed benefit as an existence proof for the encoding-advantage view of priming. Indeed, Ratcliff and McKoon (McKoon & Ratcliff, in press; Ratcliff & McKoon, in press) have recently modified their counter model to account for the both-primed benefit. The counter model now explains this effect by increasing the rate of feature extraction for studied low frequency words compared to nonstudied low frequency words. It is important to note that although enhanced discriminability as demonstrated by the both-primed benefit is a necessary condition in order to conclude priming affects low level perceptual encoding, it is not a sufficient condition (cf. Masson & McLeod, 1996). Wagenmakers, Zeelenberg, Schooler, & Raaijmakers (2000) have shown that an alternative version of the counter model can handle the both-primed benefit *without* altering the rate of feature extraction for studied words. We turn next to the REMI account of this same set of data, and show that there exists a version of this model in which a both primed benefit can be predicted without assuming perceptual enhancement due to priming. Of course there are other versions in which the both primed benefit is due to factors better termed perceptual enhancement.

REMI: A Model for Long-Term Priming

Ratcliff and McKoon (1997) argued that in order to explain the interaction of the priming effect with similarity of the alternatives one has to assume that priming changes subsequent *processing* of ambiguous information. This way, one can state that for similar alternatives, processing is of type A (i.e., 'stealing') and for dissimilar alternatives processing is of type B (i.e., 'no stealing'). However, most models of priming, such as Morton's (1969) logogen model, assume that priming affects the representation of the stimulus (e.g., by raising its level of activation). These models would seem to predict that priming effects are ubiquitous and independent of orthographic similarity of the two alternatives:

 "The main reason that other existing models cannot explain priming effects is their assumption that prior exposure to a word changes some property of the representation of the word itself. (...) When a property of the word itself changes, then processing of the word should always show facilitation relative to processing of other words." (Ratcliff & McKoon, 1997, p. 339)

The REMI model demonstrates that this claim is too strong. In the REMI model, repetition priming alters the lexical/semantic representation of the studied item by adding general contextual information to that trace. As we will explain below, REMI predicts that the impact that this stored contextual information has on the decision process varies with similarity of the alternatives: Prior study affects similar alternatives more than dissimilar alternatives.

In the REMI model, words are represented as vectors of feature values (e.g., <2, 3, 1, 4, 4,..., 4, 8>). These features can be of two kinds. First, they can contain *content*-information such as orthographic, phonological and semantic information of the particular word. Second, the features can contain *context*-information defining the setting in which the word has been encountered. Crucially, the context information is common to foils and targets in the 2-AFC task. The value of a feature defines its base rate: Common features have low values, whereas rare features have high values (sampled according to a geometric distribution). The REMI model for forced-choice perceptual identification assumes that when the alternatives are presented, their corresponding lexical memory traces are activated and matched simultaneously to the error-prone and noisy perceived features from the flash. The trace that has the greater number of matching features is chosen (guessing occurs if there is a tie). In addition, the perceived vector is influenced by the context in which the target word is flashed.

REMI models repetition priming by storing context features in the lexical representation for a word during study. The general assumption of the REM model on which REMI is based is that information is not added to existing traces if that information is already present in the trace. For well known words, content features that discriminate targets from foils are already stored in the lexical trace, and are not added. At the time of test, it is assumed that the perceived vector contains features that are unique to the current test context. Thus when the two lexical memory traces corresponding to the two alternatives are matched to the perceived vector, the studied alternative will have an advantage over a nonstudied alternative. This advantage occurs because the stored study context of the presentation will tend to match the current test context and test features, creating additional matches for the studied alternative, regardless of whether the studied alternative is the target or foil. Thus the addition of matching context features from the study episode to the lexical trace results in *bias*. The amount of bias will depend on the ratio of the number of extra matching features in comparison to the total number of features that differ between the two alternatives (i.e., the number of diagnostic features). Bias is higher when there are fewer diagnostic features, predicting the smaller bias for dissimilar alternatives. An illustration of the REMI model's predictions is given in Figures 2 and 3 (see Schooler et al., 1999, for details).

Figure 2 shows the effect of studying the *foil* alternative on a choice between similar words. An optimal decision strategy only considers diagnostic matches (i.e., matches that differentiate between the target and the foil). More specifically, the number of diagnostic matches for the target alternative is compared to the number of diagnostic matches for the foil alternative. Whichever alternative has more diagnostic matches is chosen. Therefore

performance can be determined in Figure 2 by considering the area under the curve to the right of the normative criterion of zero. In Figure 2 we see that the effect of studying the foil is to provide an additional context match to the foil alternative, causing the distribution to shift to the left. This leftward shift indicates decreased performance. Figure 3 shows the same distributions for a choice between dissimilar alternatives. When the alternatives are dissimilar there are many more diagnostic features. Since these diagnostic features are stochastically provided through perception, the potential for more diagnostic features results in greater variability in the number of diagnostic matches. As compared to Figure 2 the distributions in Figure 3 have greater variance, yet the downward shift in the distribution due to study of the foil alternative is the same. The net result is that the change in the area to the right of zero due to studying the foil is much less for dissimilar alternatives than for similar alternatives. The important point, clarified by Figure 2 and 3, is that an identical shift in the difference distributions due to prior study has more impact for similar alternatives than for dissimilar alternatives, because the variance of the difference distribution for similar alternatives is smaller than for dissimilar alternatives. Note that this model predicts bias at low flash times, because an extra match due to context occurs independent of the flash.

Thus, REMI demonstrates that priming can be explained by assuming a change in the representation of the repeated stimulus. This is a conceptually different approach than the one taken by Ratcliff & McKoon (1997). On the other hand, both models agree that priming causes bias, not an improvement in perception of the repeated stimulus.

We now turn to ways in which this simple version of the REMI model might be modified to predict the both primed benefit that has been found in recent studies (as discussed earlier). The simple version of REMI predicts (as did the 1997 version of the counter model) that prior study of both alternatives has no effect, since the context benefits are obtained by both alternatives alike, canceling each other. In one modification, the REMI model could use the approach adopted by Ratcliff and McKoon (in press), by assuming a higher rate of feature extraction for studied low frequency words: It could be assumed that study of a low frequency word has an additional effect (beyond context storage) of increasing the accuracy of the perceived vector. For example, lexical traces for low frequency words might not be complete, so that prior study could add some content features that are useful for discriminating targets from foils.

However, it is important to note that there are other approaches that can explain both primed benefits; approaches that do not posit improved perception of the flash. REMI can account for the both-primed benefit by assuming that the number of matching features needs to reach a minimum value to avoid guessing (i.e., a criterion number of matches must be exceeded). If both alternatives have too few matches the subject guesses randomly between the two alternatives (for similar discrete models using pure guessing, see for instance Anderson & Lebiere, 1998, p. 161; De Jong, 1991; Wagenmakers et al., 2000). Additional matches due to prior study serve to reduce the proportion of trials on which the subject has to guess, improving performance. To see how the model can handle frequency effects, it could be assumed that high frequency words have more common features, increasing the number of matches that occur by chance. The criterion can therefore be chosen so that high frequency words usually have enough matches to pass the criterion even without prior study, but low frequency words do not. This approach is theoretically interesting because it predicts the

both-primed benefit, as well as its dependency upon word frequency, through a criterion amount of evidence required to avert guessing.

In this modification of REMI the both-primed benefit could be viewed as a bias effect, since it is not necessary to assume an improvement in encoding of the flash. However, the assumption of a criterion and guessing when the criterion is not exceeded means that studied words are dealt with more efficiently than nonstudied words. Thus the issue of whether to label the criterion mechanism used by REMI either *bias* or enhanced *perception* is somewhat arbitrary, and might depend on that point in processing where one believes perception to cease and decision to start. This is one of those cases where the model is explicit, but it is hard to find labels that are not ambiguous.

In summary, REMI is capable of explaining various implicit priming phenomena in the 2-AFC paradigm (for detailed fits to data see Schooler et al., 1999), including the bias effect due to prior study, the interaction of prior study with similarity of alternatives, the effects of flash time, and the both-primed benefit. Both REMI and the counter model can also handle results from free response perceptual identification (in which subjects have to name the flashed word without the explicit presence of alternatives) and from the 'yes-no' task (see Schooler et al., 1999, and Ratcliff & McKoon, 1997, for details).

The critical role played by biased decision making in these priming tasks seems incontrovertible. Recent demonstrations of both-primed benefits have raised the possibility of an additional factor of perceptual enhancement. The version of the REMI model in Schooler et al. (1999) and the version of the counter model presented by Wagenmakers et al. (2000) demonstrate that such benefits could be explained by what might be considered a form of bias, although not ruling out the possibility of perceptual enhancement. Whether or not future research should demonstrate that priming sometimes produces a true perceptual benefit, as opposed to various forms of biased responding and decision making, it is worth noting that present research on word priming shows the size of priming effects due to bias are much larger than those that might be due to perceptual change.

The Present State of Long-Term Priming

Both empirically and theoretically there are at present a number of open issues in the field, in addition to the question of perceptual facilitation. One issue is raised by results of Bowers (1999). He used a paradigm almost identical to that used by Ratcliff & McKoon (1997) and found long-term repetition priming effects to be as large for dissimilar as for similar alternatives. McKoon & Ratcliff (in press) have identified one critical difference, the nature of the instructions. Differing instructions might cause differential search for memory for priming instances, and differential decision making based on the results of such memories. We hope that the near future will see quantitative models capable of handling both sets of findings in a conceptually consistent fashion.

A second issue has been raised by Masson and Bodner (2000), who suggest that implicit long-term priming effects in 2-AFC might originate through 'perceptual fluency' of the alternatives. For a choice between similar alternatives (e.g., LIED-DIED), there exists a common (in this case orthographic) ground on which the alternatives can be compared. The repeated stimulus (e.g., LIED) is perceived more fluently than its unstudied counterpart DIED, and hence it will be preferred. It is assumed that for dissimilar alternatives (e.g., LIED-SOFA) there is no common ground for comparing the two alternatives, and hence no priming effect. According to Masson and Bodner, priming effects should be observed whenever the

two alternatives share some properties on which they can be compared. Indeed, for orthographically dissimilar pairs that are semantically associated (e.g., a choice between KIWI and PEAR), Masson and Bodner *did* observe a reliable priming effect. (Of course, from the REMI perspective, such a manipulation could be viewed as producing an intermediate degree of similarity, and hence an intermediate effect size.) It should be noted that Masson and Bodner did *not* find bias effects for alternatives that were both orthographically and semantically unrelated (e.g., MOUSE- PIANO), making an explanation of their results in terms of episodic retrieval attempts induced by instructions less plausible.

ROUSE: A Model for Short-Term Priming

ROUSE (Responding Optimally with Unknown Sources of Evidence) is a model for short-term priming in the same primary task we have been considering: two-alternative perceptual word identification. In contrast with long-term priming, the prime or primes are presented a few seconds before the primary test phase. Before introducing the model, we explain the short-term priming paradigm, and present relevant data.

Figure 4 shows a typical presentation sequence (Huber et al., in press, Experiment 1). A target is briefly flashed, subsequently masked, and then followed by two alternatives, from which the subject attempts to choose the alternative that was flashed. The primes are presented immediately preceding the flashed target (instead of in a list of words studied several minutes earlier). In Huber et al.'s (in press) Experiment 1, one group of participants were told that the primes were a warning signal for the onset of the flash (the *passive* priming condition), and the primes appeared for 500 ms. Another group were instructed to make an animacy (i.e., alive or dead) decision on the primes (the *active* priming condition). The animacy decision task is indicated in Figure 4 by the dashed box, and took about three seconds to complete. (As we will see, the way subjects study the prime, actively or passively, has a profound impact on the results). The use of two primes made it possible to repetition-prime both alternatives. The alternatives used in Experiment 1 from Huber et al. (in press) were dissimilar. (Later studies in Huber et al., in press, demonstrated that similarity of alternatives does affect the results, in complex ways that are predicted by ROUSE, and the reader is referred to that article for details.)

Although Huber et al. (in press) used many types of priming relations, we will discuss here only the results for repetition priming. There were four conditions: (1) neither the target nor the foil were primed, (2) both the target and the foil were primed, (3) the target was primed, and (4) the foil was primed. The results can be seen in Figure 5. Remarkably, when both alternatives were primed by presenting them immediately before the target flash, performance is decreased compared to when neither alternative was primed, a result obtained for both passive priming and active priming (see the first two columns of Figure 5). Thus, whereas in long-term priming a both-primed benefit is observed, short-term priming is characterized by a 'both-primed *deficit*'. The second surprising result is that subjects prefer the primed alternative in the passive condition (similar to the bias findings in long-term priming), but this preference is *reversed* in the active priming condition (see the last two columns in Figure 5). In other words, after thorough processing of the primes, subjects have a tendency *not* to choose the primed alternative but rather to choose the unprimed alternative. Next, we will introduce the ROUSE model capable of handling these results (as well as many others results that space considerations prevent us from discussing).

ROUSE models the primary task in essentially the same fashion as REMI: The perceived information is matched to the two alternatives, producing vectors of matches and mismatches. More precisely the results of matching a given feature are represented by one of two states: 'match' or 'mismatch', termed 'ON' and 'OFF' for short. Prior to the flash, all features are 'OFF'. However, three sources of activation, operative during task execution, can change the state of any feature to 'ON'. Figure 6 illustrates these three sources of activation: First, features from the target alternative (e.g., LIED) can be activated by perceived features extracted from the flash (e.g., LI) with probability β. This probability β indicates perceptual encoding and increases with flash time. Second, noise, either due to perturbation from the mask or from random fluctuations in the perceptual system itself, turns features 'ON' randomly in the target vector and the foil vector with probability γ. Finally, features that are part of the *primes* can activate corresponding features in either alternative with probability α. This source confusion results in a bias for alternatives similar (or identical) to prime words. After the three sources of activation have turned features 'ON' and left other features 'OFF', a near-optimal decision is made between the two alternatives: Likelihoods are calculated for each alternative, and the higher likelihood determines the choice. The critical feature of ROUSE occurs during the likelihood calculations: These take into account the possibility that features may be turned on by primes. In particular, the evidence provided by a feature that might have arisen from a prime is discounted. This discounting of features known to be in a prime word can result in a reversal of the bias effect.

Consider a diagnostic feature that has been turned 'ON' in alternative A (e.g., the letter L in LIED when the choice is between LIED and DIED). This feature provides strong evidence in favor of LIED because it could have been turned on only by correct perception or by noise (and γ in ROUSE is actually a very small probability). However, suppose that LIED had been presented as a prime. Now the L feature being 'ON' provides less evidence because it could have been turned 'ON' by the prime, rather than correct perception or noise.

In an optimal system, the likelihood calculations would use an accurate estimate for the probability of source confusion due to the primes (i.e., the estimate of α denoted α'). However, it is assumed that passive priming leads the estimate to be slightly too low, producing too little discounting, and hence a preference for the alternative that has been primed (since the 'ON' features will be produce too much evidence for the primed alternative). It is assumed that active priming leads the estimate to be slightly too high, producing the opposite outcome--a preference for the alternative that was not primed. These mechanisms illustrate how ROUSE predicts the bias effects when one alternative is primed.

Consider next what happens when both alternatives are primed. By symmetry, neither alternative gains an advantage. However, the primes tend to turn 'ON' extra features in both alternatives. These extra 'ON' features produce extra noise, tending to drown out the signal from the features turned on by the flash. This extra noise is what produces the both-primed deficit that is observed in both priming conditions.

At this point it is worth noting that ROUSE models short-term priming as a bias or preference effect.[2] ROUSE does not assume priming to result in a higher rate of feature extraction for the repeated stimulus (i.e., a perceptual effect). However, Huber et al. (in press) did find a small but reliable benefit of associative priming when a prime word (e.g., MISTAKE) was associatively related to both choice words (e.g., ERROR and WRONG)

compared to when the prime was not related to the choice words (see also Masson & Borowsky, 1998). Since the semantic overlap between prime and target is 100% for repetition priming, it is feasible that there might in fact be a semantic both-primed *benefit* in short-term repetition priming, albeit overshadowed by a much larger orthographic both-primed *deficit*. At present, ROUSE does not provide an explanation for the small both-primed benefit occasionally observed when the prime is associatively related to the choice words. Extensions of the model to handle such findings await further research, as does any final judgment whether such findings demonstrate perceptual enhancement rather than some more subtle form of bias.

Space does not permit discussion of the way in which ROUSE predicts results for other sorts of prime relations. In particular, we cannot discuss the many non-intuitive predictions that ROUSE makes as one manipulates the similarity of primes to choices, and choices to each other. Nevertheless, the fact that these predictions are borne out in the results is a major source of evidence in favor of the model (see Huber et al., in press, for details)

Discussion of ROUSE

The general notion that people sometimes mistrust salient information is not new. For instance, social psychologists have found a reversal of the priming effect for evaluative judgements (Lombardi, Higgins, & Bargh, 1987): Participants able to recall the priming event gave evaluative judgements *opposite* to the prime, whereas participants unable to recall the priming event tended to give evaluative judgements *consistent* with the prime. However, this general notion of discounting had not been applied to visual word recognition or formulated in a process model prior to the development of ROUSE. In addition, several priming phenomena can be re-interpreted when one considers how ROUSE can be applied to those phenomena. First, consider sub-threshold ('subliminal') priming: When a prime is presented so briefly that the subject is unable to identify it, effects of priming are still observed and can be larger than for supra-threshold priming. For example, Murphy and Zajonc (1993) had subjects give liking ratings to Chinese ideographs. These ideographs were preceded by positive or negative affective primes, shown either for a very brief duration or for a longer duration. For very brief durations, the priming effect was congruent (i.e., a positive prime led to a positive evaluation for the ideograph and vice versa). However, this effect disappeared and even tended to reverse when the affective primes were presented for longer durations. Murphy and Zajonc appeal to a theoretical relation between consciousness, emotion, and cognition to account for these interesting results. We believe ROUSE can give a more parsimonious explanation of their data: When subjects cannot identify the prime, they may not discount prime information appropriately. In fact, subjects might not be discounting information at all (i.e., their estimate of α may equal zero) causing a preference for a prime-related response. In contrast, when subjects can clearly see the prime, they may engage in a discounting process that suppresses prime-related responding. If subjects discount prime information too much (i.e., their estimate of α is greater than α), this might even lead to a preference for prime-unrelated responding.

ROUSE can provide an alternative explanation concerning the effects of orthographic priming in visual word recognition in a similar fashion.[3] When an orthographically related prime (e.g., HEAT) is presented immediately before presentation of the target (e.g., MEAT), this sometimes leads to inhibition and sometimes to facilitation of the response to the target relative to when an orthographically unrelated prime (e.g., TOSS) is presented. More

specifically, in what is known as 'form-priming', the rapid successive presentation of a masked orthographically related prime and the target increases performance (e.g., Evett & Humphreys, 1981). In form priming, subjects are usually unable to identify the prime and perceive prime and target as a single perceptual event. On the other hand, when the prime is clearly presented for a relatively long duration preceding the onset of the target, inhibitory effects due to priming the target with an orthographically similar word (i.e., a so-called 'neighbor') are sometimes observed (e.g., Colombo, 1986). The form-priming results are commonly interpreted as being due to *perceptual facilitation*, whereas the inhibitory neighbor priming results are often attributed to some kind of lexical suppression through inhibitory connections between simultaneously activated words. In ROUSE, facilitatory form-priming can arise once subjects do not discount prime-related information enough, and this is best characterized as a *preference* or bias effect. Further, ROUSE can account for the phenomenon of inhibitory neighbor priming by assuming subjects sometimes discount prime-related information too much. Rather than being due to lexical suppression, inhibitory neighbor priming can thus be due to a source monitoring problem (see O'Seaghdha & Marin, 2000, for a related account).

In general, ROUSE shows that the way subjects study the primes (i.e., passively or actively) can have a profound influence on the magnitude and even the direction of the priming effect. When one considers the amount of attention paid to the primes is sometimes not constrained or manipulated, it comes as no surprise that, for example in semantic priming, results are occasionally contradictory and ambiguous (e.g., Neely 1991).

General Discussion

In this chapter we have presented a common model framework for the task of threshold word identification assessed by two-alternative forced-choice testing. The models are based on an optimal approach to decision making utilizing Bayesian principles. Using this model framework we have shown how it is possible to model both long-term and short-term priming effects, thereby demonstrating how a variety of phenomena of implicit memory can be placed in a common conceptual framework in a rigorous, quantitative manner. Of course, the mechanisms of priming are somewhat different in the two cases, as one might expect.

First, REMI is a model for long-term priming that gives an explanation of priming based upon the storage of context information in the lexical trace during study of the prime. Because the context during the study phase will tend to match the context present in the test phase, a system using optimal decision principles will tend to choose the studied alternative to the nonstudied alternative. The greater number of diagnostic features for dissimilar choices produces a great deal of variability due to chance, reducing the impact of extra matching features due to priming (see Figures 2 and 3). REMI is also able to account for the both-primed benefit that is found for low frequency words by setting a criterion that has to be exceeded in order to avoid guessing. This demonstrates that the both-primed benefit does not need to imply a higher rate of feature extraction for repeated words.

Second, ROUSE is a model for short-term priming that gives an explanation of priming based upon source confusion for prime activated features. In producing an optimal decision, features that might have been activated by the prime provide a discounted level of evidence. The discounting mechanism gives a parsimonious account of many seemingly contradictory data (e.g., the reversal of the priming effect with increased processing of the prime). In addition, priming itself produces chance variability, accounting for the both-primed

deficits that are observed with short-term repetition priming (see the first two columns of Figure 5).

REMI and ROUSE both account for priming through what might be considered *bias*, both models in their simplest form assuming the effect of prior study does not act to change the quality of information that is entering the system, but rather to change the *impact* that the incoming information has on the decision process. Generally speaking, the effect of priming is to tune the decisional mechanisms toward likely events, that is, events that recently occurred in the same environment. REMI and ROUSE are based on Bayesian principles and the idea that decisions are made to optimize performance in the face of noisy data. Although the brain is probably not calculating log odds or likelihood ratios, it might have learned to approximate such a calculation over developmental time in order to behave in a near-optimal fashion when confronted with a noisy and uncertain environment. One of the advantages of using a Bayesian approach is that it provides a rational framework for modeling that is not ad hoc.

We believe that priming phenomena, and implicit memory phenomena in general, can be better understood when they are modeled in a consistent way. We have demonstrated that both REMI and ROUSE, using a common theoretical framework, are capable of handling some of the most important extant findings in the field of priming as well as recent new data (e.g., the findings of the both-primed benefit in long-term memory, the both-primed deficit in short-term memory, and the reversal of preference when the nature of prime processing is changed). Much of this research is still at an early stage, but the successes we have presented in this chapter hold out hope for rapid theoretical progress in the near future.

References

Anderson, J. R., & Lebiere, C (1998). *The atomic components of thought*. Lawrence-Erlbaum Associates.

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

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

Colombo, L. (1986). Activation and inhibition with orthographically similar words. *Journal of Experimental Psychology: Human Perception and Performance*, 12, 226-234.

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

Church, B. A., & Schacter, D. L. (1994). Perceptual specificity of auditory priming: Implicit memory for voice intonation and fundamental frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 521-533.

Evett, L. J., & Humphreys, G. W. (1981). The use of abstract graphemic information in lexical access. *Quarterly Journal of Experimental Psychology*, 33, 325-350.

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, 91, 1-67.

Graf, P., Squire, L. R., & Mandler, G. (1984). The information that amnesic patients do not forget. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 164-178.

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

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

Lombardi, W. J., Higgins, E. T., & Bargh, J. A. (1987). The role of consciousness in priming effects on categorization: Assimilation versus contrast as a function of awareness of the priming task. *Personality and Social Psychology Bulletin*, 13, 411-429.

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, Wyoming.

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., & Freedman, L. (1990). Fluent identification of repeated words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 355-373.

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.

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

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

Murphy, S. T., & Zajonc, R. B. (1993). Affect, cognition, and awareness: Affective priming with optimal and suboptimal stimulus exposures. *Journal of Personality & Social Psychology*, 64, 723-739.

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

O'Seaghdha, P., & Marin, J. W. (2000). Phonological competition and cooperation in form-related priming: Sequential and nonsequential processes in word production. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 57-73.

Raaijmakers, J. G. W., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review*, 88, 93-134.

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

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. (in press). Modeling the effects of repetition and word frequency in perceptual identification. *Psychonomic Bulletin & Review*.

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

Roediger, H. L., III, Weldon, M. S., Stadler, M. L., & Riegler, G. L. (1992). Direct comparison of two implicit memory tests: Word fragment and word stem completion. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 1251-1269.

Salasoo, A., Shiffrin, R. M., & Feustel, T. C. (1985). Building permanent codes: Codification and repetition effects in word identification. *Journal of Experimental Psychology: General, 114,* 50-77.

Schacter, D. L. (1990). Priming and multiple memory systems: Perceptual mechanisms of implicit memory. In D. L. Schacter & E. Tulving (Eds.), Memory systems 1994 (pp. 233-268). Cambridge, MA: MIT Press.

Schooler, L. J., Shiffrin, R. M., & Raaijmakers, J. G. W. (1999). Theoretical note: A model for implicit effects in perceptual identification. Manuscript submitted for publication.

Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM—retrieving effectively from memory. *Psychonomic Bulletin & Review*, 4, 145-166.

Shimamura, A. P., & Squire, L. R. (1984). Paired-associate learning and priming effects in amnesia: A neuropsychological study. *Journal of Experimental Psychology: General, 113*, 556-570.

Squire, L. R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychological Review, 99*, 195-231.

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

Wagenmakers, E. M., Zeelenberg, R., Schooler, L. J., & Raaijmakers, J. G. W. (2000). A criterion-shift model for enhanced discriminability in perceptual identification: A note on the counter model. (manuscript submitted for publication).

Zeelenberg, R., Shiffrin, R. M., & Raaijmakers, J. G. W. (1999). Priming in a free association task as a function of association directionality. *Memory & Cognition, 27,* 956-961.

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

Footnotes

[1] In earlier work, we used the more general term 'enhanced discriminability', denoting that the system is better in distinguishing the target from the foil. In the following, we use the term 'both-primed benefit' as a shorthand for 'enhanced discriminability due to priming of both alternatives'.

[2] Huber et al. (in press) use the label 'preference' instead of 'bias', because 'bias' is, according to their view, a theory-laden concept that might suggest, wrongly, signal-detection theory to be applicable.

[3] A number of variations in orthographic priming were explored by Huber et al. (in press) within a forced choice paradigm, and were fit quite successfully by ROUSE. Prior research using other paradigms is the subject of this section.

Figure Captions

Figure 1. Percentage correct identifications in forced-choice auditory word identification as a function of study condition (Zeelenberg, Wagenmakers, & Raaijmakers, 2000, Experiment 2).
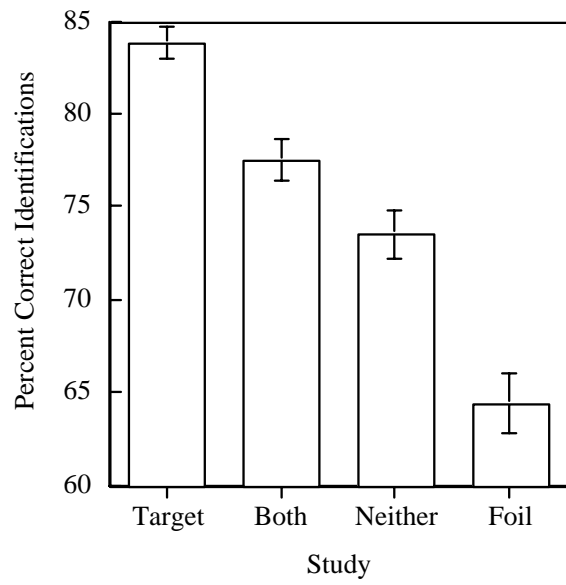
Figure 2. The distributions for the difference between the number of matches for the target and a *similar* foil (e.g., LIED- DIED), both for a studied foil as for an unstudied foil.
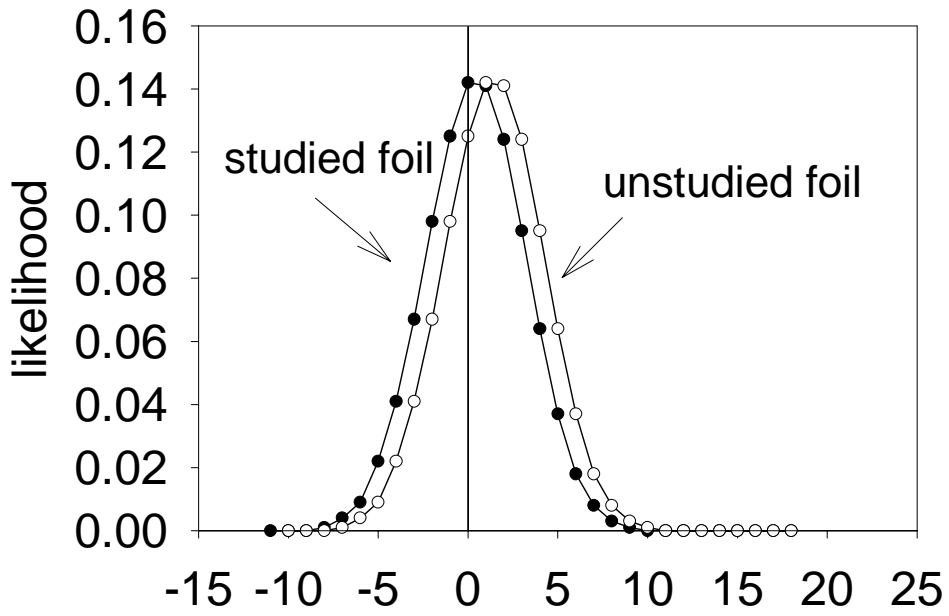
Figure 3. The distributions for the difference between the number of matches for the target and a *dissimilar* foil (e.g., LIED-SOFA), both for a studied foil as for an unstudied foil.

Figure 4. A typical sequence of displays in a 2-AFC short-term priming experiment. The display contained within the dashed box only appears in the active priming condition.

Figure 5. Results and ROUSE-predictions for data from Huber et al., in press (Experiment 1). Error bars are two standard errors of the mean.

Figure 6. The ROUSE model for short-term priming.

# of target diagnostic matches - # of foil diagnostic matches

# of target diagnostic matches - # of foil diagnostic matches

PRIME1

PRIME2

Active Priming Task

PRIME2

PRIME1

Prime: 500ms

TARGET

Target: 17,25,..,117ms
(set individually)

Mask: 500ms
- target duration

TARGET   XFOILX

Presentation
Sequence

Test display: until response