Journal of Experimental Psychology:
Human Perception and Performance
2005, Vol. 31, No. 1, 40–61

0096-1523/05/$12.00 DOI: 10.1037/0096-1523.31.1.40

Confusion and Compensation in Visual Perception: Effects of Spatiotemporal Proximity and Selective Attention

Christoph T. Weidemann
Indiana University Bloomington

David E. Huber
University of Maryland

Richard M. Shiffrin
Indiana University Bloomington

The authors investigated spatial, temporal, and attentional manipulations in a short-term repetition priming paradigm. Brief primes produced a strong preference to choose the primed alternative, whereas long primes had the opposite effect. However, a 2nd brief presentation of a long prime produced a preference for the primed word despite the long total prime duration. These surprising results are explained by a computational model that posits the offsetting components of source confusion (prime features are confused with target features) and discounting (evidence from primed features is discounted). The authors obtained compelling evidence for these components by showing how they can cooperate or compete through different manipulations of prime salience. The model allows for dissociations between prime salience and the magnitude of priming, thereby providing a unified account of "subliminal" and "supraliminal" priming.

Information processing is influenced by the prior presentation of related stimuli (i.e., primes), even if these stimuli are irrelevant to the task at hand. Such priming effects are used to investigate knowledge representation, the dynamics of perception and retrieval, and decisional factors as determined by task demands (e.g., Pecher, Zeelenberg, & Raaijmakers, 1998; Tipper, 2001).

In examining priming phenomena, it is useful to differentiate between paradigms that separate prime and target by a substantial amount of time (i.e., long-term priming) and those that present primes immediately prior to targets (i.e., short-term priming). Unless otherwise specified, the term *priming* in this article refers to short-term priming. Priming has been studied for tasks such as naming, identification at threshold, and lexical decision. In this study, we investigated priming in forced choice perceptual identification of briefly flashed words, which we have found to be a useful method for separating perceptual and decisional factors.[1]

Many studies have used priming as a tool for studying the structure of the mental lexicon. For instance, a lexical decision response to a word (e.g., *BUTTER*) is faster when primed by a semantically related word (e.g., *BREAD*) than when primed by a less related control word (e.g., *DOCTOR*; Meyer & Schvaneveldt, 1976). It is somewhat surprising that such priming occurs even when the primes themselves are extremely difficult to identify. In fact, priming occurs even when the observer cannot judge whether any prime has been presented (e.g., Marcel, 1983). Such effects are often referred to as showing *subliminal* or *unconscious* perception (for reviews, see Merikle & Daneman, 1998; Merikle, Smilek, & Eastwood, 2001).[2]

Christoph T. Weidemann and Richard M. Shiffrin, Department of Psychology, Indiana University Bloomington; David E. Huber, Department of Psychology, University of Maryland.

Parts of this article are based on Christoph T. Weidemann's diploma thesis at the University of Bonn, Bonn, Germany. Portions of this work were presented at the 35th Annual Meeting of the Society for Mathematical Psychology, Miami, Ohio, July 2002; the First and Second Annual Summer Interdisciplinary Conferences, Squamish, British Columbia, Canada, July–August 2002 and 2003, respectively; the Third Annual Summer Interdisciplinary Conference, Cavalese, Italy, June–July, 2004; the 33rd European Conference on Mathematical Psychology, Bremen, Germany, August 2002; and the 43rd Annual Meeting of the Psychonomic Society, Kansas City, Missouri, November 2002.

This research was supported by National Institute of Mental Health Grants MH63993 and MH12717. We thank Andrew L. Cohen for his help in resolving some programming issues; Jamila Siddiqui for her assistance in administering Experiment 3; and Arndt Bröder, Shane T. Mueller, Jörg Rieskamp, and, Adam N. Sanborn for helpful comments on earlier versions of this article.

Correspondence concerning this article should be addressed to Christoph T. Weidemann, Department of Psychology, Indiana University, 1101 East 10th Street, Bloomington, IN 47405-7007. E-mail: cweidema@indiana.edu

[1] In this article, we use the terms *decision* and *decisional* in a very general sense to mean some aspect of evidence evaluation that is dictated by the pressures of optimal performance. Note that this does not imply that the decision process is subject to conscious awareness or that it always yields an optimal outcome. It is important to note that throughout this article, our classification of effects as *perceptual* (i.e., influenced by the perceptual response to the target presentation) and *decisional* or *preferential* (i.e., independent of the perceptual response to the target presentation) is somewhat different from the concepts of *sensitivity* and *bias* used in signal detection theory (e.g., Macmillan & Creelman, 1991) and does not imply differential stages or levels of processing (e.g., preferential effects need not be postperceptual; see Huber, Shiffrin, Lyle, & Ruys, 2001, for a detailed discussion of this terminology).

[2] The question of when (if ever) perception should be considered to occur unconsciously is highly dependent on a theory of consciousness. Such a theory, however, is seldom explicitly specified, nor is such an endeavor necessary to investigate the effect. Thus, for the sake of clarity, we avoid the terms *subliminal* and *unconscious* and instead simply describe these results as showing a dissociation between priming and prime detection (prime detection is indexed by above-chance performance in explicitly identifying trials containing a prime).

Priming Without Prime Detection

Dissociations between priming and prime detection are usually achieved by either presenting the prime(s) very briefly (usually followed by a backward mask) or by using a task that diverts attention from the prime(s) (see Merikle et al., 2001). Naive intuition suggests that the amount of priming should be directly related to the strength of perception: A highly salient prime should cause more priming than a minimally salient (e.g., undetected) prime. Such a prediction is consistent with theories in which activation spreads from the prime's lexical node to other lexical nodes that are semantically related (see, e.g., McNamara, 1992, for a review of these mechanisms and some specific theories of priming). However, this prediction is often violated. For instance, Marcel (1983, Experiment 5) found sizable priming for undetected primes, an effect that increased when the primes where repeated on different trials, even though the primes remained undetected, and Forster and Veres (1998) found robust form priming for word targets with briefly presented (i.e., 50-ms, pre- and postmasked) primes, which was disrupted when the primes were displayed for 500 ms. Furthermore, Huber (2005b; see also Figure 8 of Huber & O'Reilly, 2003) found that the priming effect for very brief prime presentations (as short as 17 ms, postmasked by the target presentation) could be as large or larger than that for prime durations that were clearly above threshold. This dissociation between priming magnitude and prime detection is poorly accounted for by most theories of priming due to their constraint to produce a direct relationship between the magnitude of priming and the degree of prime salience.

Early theories of priming can be separated into activation-based (Collins & Loftus, 1975; McNamara, 1992) and memory retrieval accounts (Dosher & Rosedale, 1989; Ratcliff & McKoon, 1988). In either case, the prime directly adds to target processing, either as a function of residual activation or through enhanced memory cuing. Therefore, both model classes suppose a positive relationship between the availability or strength of the prime and the degree of priming benefit. This direct relationship between prime salience and priming magnitude is true even of newer, linguistically sophisticated models such as the DRC (dual-route cascaded) model of visual word recognition and reading aloud (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001), the MROM (multiple read-out model; Grainger & Jacobs, 1996), and connectionist accounts (as exemplified by Plaut & Booth, 2000). Although these theories are tremendously successful in explaining a wide variety of lexical phenomena, they do not contain mechanisms for capturing this fundamental dissociation between prime salience and the strength of priming. It is not our intent to supplant such theories but, rather, to suggest that they require augmentation with compensatory mechanisms to capture this dissociation, particularly in the case of above-threshold presentations: The severity of the dissociation is seen most clearly in recent research by Huber et al. (2001) in which increasing prime salience to an excessively high degree (as compared with traditional priming paradigms) not only failed to further increase priming benefits but actually resulted in robust priming deficits.

To explain these and many other unanticipated priming effects, Huber et al. (2001) developed a new model of short-term priming called *responding optimally with unknown sources of evidence* (ROUSE). The critical effects were found in two conditions, termed *passive priming* and *active priming*. Passive priming involved an attempt to reduce prime saliency and attention allocated to the prime by instructing observers to ignore primes (which were presented for 500 ms; Huber et al., 2001) and to treat them merely as a warning signal for the subsequent target flash. Active priming involved an attempt to maximize prime salience either by asking observers to make judgments about prime words (e.g., "Could this word be a verb?"; Huber et al., 2001) or by presenting primes for longer than usual durations (Huber, Shiffrin, Lyle, & Quach, 2002). The results in these forced choice perceptual identification tasks were initially unanticipated: Although passive priming produced a benefit when the target was primed and a deficit when the incorrect alternative (the *foil*) was primed (i.e., a preference for the primed word), active priming produced the opposite pattern of results, with a deficit when the target was primed and a benefit when the foil was primed (i.e., a preference against the primed word). Thus, depending on the prime condition (i.e., priming of target or foil) and the nature of prime processing (i.e., passive or active), the same prime word could cause an increase or a decrease in performance.

The ROUSE model accounts for these (and other) findings by incorporating two key assumptions:

1. Prime and target features become confused into one percept (*source confusion*).

2. The expectation of source confusion is used to compensate the evidence assigned to primed items (*discounting*).

The details are given below, but the key result is that the direction and amount of priming are determined by the difference between the degree of confusion caused by primes and the estimate of source confusion when the system calculates evidence. With two separate factors determining this difference, the magnitude of the difference, and therefore the magnitude of priming, need not increase as a function of prime saliency. In fact, Huber et al. (2001) found that making the difference positive for passive priming (low salience) and negative for active priming (high salience) accounted for the observed reversal in the direction of priming.

Within this framework, it is easy to explain the dissociation between prime detectability and degree of priming: If the difference between source confusion and discounting is about equal for primes near detection threshold and primes above threshold, then the degree of priming will not differ between these conditions. Why might this difference be roughly equal in these two cases? For a prime just below detection threshold, it seems unlikely that the decision system could know that source confusion has taken place (i.e., there is little or no discounting). Of course, for such a weak prime, there will be less source confusion, but the full force of this source confusion will be felt in the absence of discounting, and the net result will be a positive priming effect. For above-threshold primes, the size of the difference (and the direction of the difference) will be determined by the estimate of source confusion. Accurate estimation will produce a zero difference, but if the estimate is slightly in error, the difference will again be small and not necessarily larger than that for near-threshold primes.

Spatiotemporal Determinants of Priming

It is a key assumption of the ROUSE model that features of the prime become confused with the target percept. But what causes or allows such confusions? Plausible candidates are

1. orthographic, phonological, and semantic similarity of prime and target;
2. nearby spatial position of prime and target; and
3. nearby temporal position of prime and target.

Not only prime saliency and detectability but also orthographic, phonological, and semantic similarity have received considerable empirical and theoretical attention. However, effects of spatiotemporal separation between prime and target have been the focus of comparatively few studies.

Temporal Relations

Humphreys, Besner, and Quinlan (1988) used repetition priming in a perceptual identification task. They compared identification performance for targets that were preceded by a brief (<55-ms, premasked) prime with that for targets preceded by a longer (300-ms, no-premask) prime.[3] Brief primes increased and long primes decreased identification accuracy. In their Experiment 4, Humphreys et al. inserted a 120-ms pattern mask between a 180-ms (no-premask) prime and the target and found positive priming. These results align well with recent results that inspired and tested the ROUSE model (e.g., Huber et al., 2001; Huber, Shiffrin, Quach, & Lyle, 2002) and with the results presented in the current article. However, because temporal separation between prime and target was confounded with prime duration in Humphreys et al.'s experiments, it is unclear whether (a) priming was affected by temporal separation or (b) these results can be entirely attributed to the effect of prime duration.

Hochhaus and Marohn (1991) also varied the timing between prime and target. Their Experiment 1 used a perceptual identification task: A 500-ms prime was followed by the target (with interstimulus intervals [ISIs] of 0, 250, or 1,500 ms). Using a single test word and same–different testing, Hochhaus and Marohn found decreased sensitivity for primed targets at the 0-ms ISI and no priming effects for longer ISIs (they interpreted the former result as a form of repetition blindness, a phenomenon we take up in this article's General Discussion).[4] This result suggests that priming is disrupted or diminished when primes and targets are separated by a sufficiently long (e.g., 250-ms) intervening stimulus. Therefore, to better understand the temporal dynamics of priming, in Experiment 1 we separated primes and targets by much briefer intervals than those used by Hochhaus and Marohn.

Spatial Relations

In their Experiment 4, Hochhaus and Marohn (1991) manipulated the vertical distance of a 500-ms prime relative to a central target. They found a substantial priming effect at all tested distances (between 0° and 1.68° of visual angle), but the effect was larger for smaller distances. However, this finding of a decreasing priming effect with increasing distance between prime and target is not universal. Stankiewicz and Hummel (2002), for example, demonstrated that visual object priming is not affected by a distance of 5° of visual angle between prime and target. Dill and Edelman (2001) offered a possible resolution to such a discrepancy by suggesting that translation-invariant priming only applies to local features (and objects that can be identified on the basis of only such features), whereas priming for global structure depends on spatial proximity. Ashby, Prinzmetal, Ivry, and Maddox (1996), using a different paradigm, showed that errors in feature binding (e.g., illusory conjunctions) depend on the spatial proximity of the features to each other. This result supports the claim that spatial confusions may be partly responsible for priming when feature binding is necessary for the task.

In the experiments of Huber et al. (2001), the primes, target flash, and choices were arrayed in different spatial positions (primes up and down, target central, choices left and right, with less than 3° of visual separation). Sizable priming effects were observed (albeit in a direction that changed with priming condition). In more recent experiments (reviewed by Huber & O'Reilly, 2003), a single prime word was presented twice simultaneously—one instance above the other, just touching—followed by a central target that overlapped the lower half of the upper prime and the upper half of the lower prime. These experiments yielded considerably larger priming effects than had previous studies, but it is likely that the increased priming was due less to the increased proximity of primes than to the fact that there was only a single, unique prime word. Such a conclusion is suggested by the results of Huber et al.'s (2001) Experiment 4, in which a single prime was repeated but without spatial overlap with the target: Enhanced priming was observed in this condition, too.

The spatial manipulations discussed above occurred across different experiments and priming conditions. Therefore, we designed the present experiments to provide better data concerning the role of spatial distance in priming. Furthermore, we were interested in possible interactions between spatial and temporal distance manipulations, so we covaried these factors in Experiment 1.

---

[3] Humphreys et al. (1988) labeled these conditions as involving *masked* and *unmasked* primes, respectively (referring to the presence of a forward mask in the former case and its absence in the latter). Because recent evidence suggests that the crucial difference between the two priming conditions was not the presence or absence of a forward mask but, rather, prime duration (e.g., Huber, Shiffrin, Quach, & Lyle, 2002; cf. also the results of Humphreys et al.'s, 1988, Experiment 4), and because primes were postmasked by the target presentation in both conditions, we prefer to refer to these prime presentations as *brief* and *long*, respectively.

[4] Hochhaus and Marohn (1991, Experiment 1) also used a 0-ms ISI condition with a 250-ms prime. This shorter prime duration produced no repetition blindness (in fact, nonsignificant positive priming was found). Other studies have produced significant positive priming for prime durations as long as 500 ms (e.g., Huber et al., 2001). Taking into account differences in instructions, prime presentation, and task, which can have profound effects even when prime duration is held constant (see, e.g., the current Experiment 3), the qualitative difference in repetition blindness for the two prime durations is naturally predicted by the ROUSE model.

Combining Primes Across Time and Space

A related and perhaps more important issue concerns the way that multiple primes combine at different times and different positions to influence target processing. This is a critical issue because ROUSE posits two factors that determine priming: source confusion and discounting. These separate factors may be differentially affected by spatial and temporal variations, providing converging evidence for the existence of these factors. In Experiments 1 and 2, we examined the effect of multiple prime presentations of different durations and spatial positions in the same trial. We generalized these results in Experiment 3, presenting two primes concurrently but with instructions to ignore one and actively process the other.

Perceptual Identification

Most priming studies use tasks like lexical decision or naming, with response time as the dependent measure. It has long been recognized that interpretation of such findings is difficult because effects could be due to perceptual changes, changes in the preference to respond with a primed alternative (also known as *decisional bias*; see Footnote 1), or both. To distinguish perception from preference, Huber et al. (2001) borrowed a technique used with great success by Ratcliff and McKoon (1997; see also Ratcliff, McKoon, & Verwoerd, 1989) for the study of long-term priming. They used a two-alternative forced choice (2-AFC) perceptual identification paradigm requiring the matching of a briefly flashed and postmasked target word to one of two subsequently presented choice words (accuracy was the measure of interest). In Huber et al.'s (2001) experiments, the presentation of the target was closely preceded by the presentation of two prime words. Perceptual and preferential priming factors can be separately assessed by testing conditions that prime both choices, only the correct choice (i.e., the target), only the incorrect choice (i.e., the foil), or neither choice. A difference between the foil-primed condition and the neither-primed condition indicates a preference effect, because the display in both conditions is identical until the presentation of the choice words (the same is true for a comparison of the both-primed condition and the target-primed condition). Conversely, a benefit for the both-primed condition over the neither-primed condition indicates a perceptual effect, because the target has been primed, but with the foil also primed, decisional effects are properly controlled. Studies using this paradigm have varied the relation of primes to choice words, the similarity of primes to choice words, the duration of primes, attention to primes, the similarity of primes to each other, and the similarity of choice words to each other, and they have related memory for primes to the pattern of priming (Huber, 2005b; Huber, Shiffrin, Lyle, & Quach, 2002; Huber et al., 2001; Huber, Shiffrin, Quach, & Lyle, 2002).

A closely related method for separating perceptual and preferential factors is the use of a single test word with a same–different task (e.g., Hochhaus & Johnston, 1996; Paap, Johansen, Chun, & Vonnahme, 2000, Experiments 4a and 4b). A signal detection analysis of such data, applied separately for primed test words and unprimed test words, produces a bias measure that indicates preference effects (i.e., did the bias change for primed words?) and a sensitivity measure that indicates perceptual effects (i.e., after bias was controlled for, did performance increase?). The results of both types of studies highlight the importance of the decision process in short-term priming: Preference effects were found to be large, and they changed direction depending on the way primes were processed, whereas perceptual effects were much smaller or missing (e.g., Huber et al., 2001; Paap et al., 2000, Experiments 4a and 4b).

ROUSE

To account for a variety of findings from 2-AFC perceptual identification paradigms, Huber et al. (2001) developed a quantitative model of short-term priming, ROUSE. This model incorporates the offsetting mechanisms of source confusion and discounting. ROUSE is a Bayesian decision theory, similar in spirit to the REM (retrieving effectively from memory) model of Shiffrin and Steyvers (1997; see also Shiffrin & Steyvers, 1998) and its recent variants (e.g., Schooler, Shiffrin, & Raaijmakers's, 2001, REMI [retrieving effectively from memory, implicit] model for long-term priming). In ROUSE, all representations are feature based, and a Bayesian decision process arrives at a nearly optimal response conditioned on noisy perception and/or noisy memory. The theory has two parts: *Unknown sources of evidence* refers to the assumption that features of the choice words can be activated by the prime presentation, the target presentation, or visual noise, with the actual source of activation being unknown to the system (i.e., source confusion). In isolation, this factor causes a preference for prime-related choice words. The Bayesian decision process (*responding optimally*) removes, or even reverses, this preference through the discounting of evidence from features known to have been in the primes. The reversal of preference across experimental conditions is explained in ROUSE terms by too little discounting for brief prime durations and appropriate or too much discounting for longer prime durations (Huber et al., 2001; Huber, Shiffrin, Quach, & Lyle, 2002). Below, we present the *unknown sources of evidence* part of the theory first, and then we review the Bayesian decision process.

Unknown Sources of Evidence

In ROUSE, each choice word is represented by a vector of binary features (typically set to 20; Huber et al., 2001). Features in the choice words can be independently activated (i.e., perceived) by three sources of evidence. Assume for now that the two choice words share no features. See Figure 1 for an illustration of the following activation processes (Huber et al., 2001):

1. *Prime presentation:* With probability α, primed features are activated.

2. *Target presentation:* With probability β, target features are activated.

3. *Visual noise:* With probability γ, any feature is activated.

Source confusion, if not somehow countered, will always lead to a preference for primed choice words due to the additional (often misleading) evidence supplied by the prime. The probability of the prime activating primed features of a choice word determines the amount of additional evidence, and therefore, α is the primary determinant of source confusion. However, the decision process in ROUSE counteracts this preference by differentially weighting evidence on the basis of its diagnostic value.

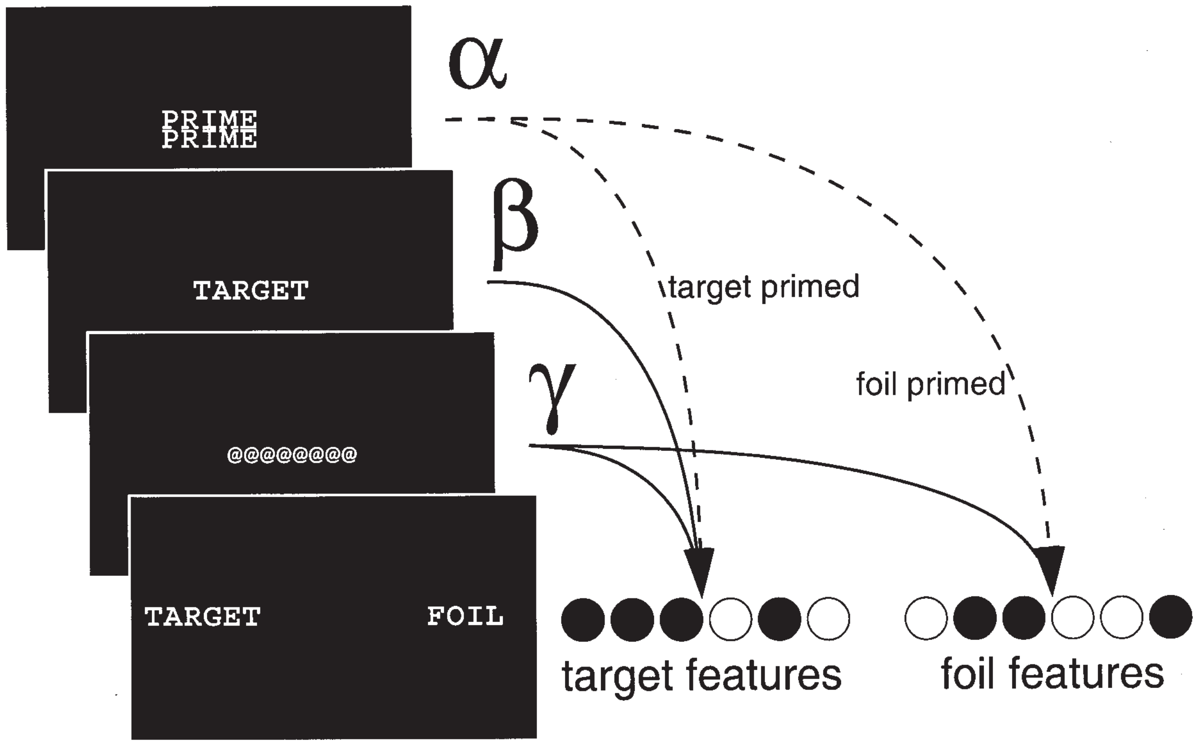

Figure 1. In ROUSE (responding optimally with unknown sources of evidence), words are represented as vectors of binary features. Choice-word features can be independently activated by three sources of evidence: the prime presentation (with probability α), the target presentation (with probability β), and visual noise (with probability γ). Feature activation by the prime presentation depends on the similarity of the prime to the choice words, which is assumed to be 0 for unrelated primes and 1 for identical primes (denoted by the dashed arrows). The prime is repeated in the figure solely for consistency with the prime presentation in the present experiments.

Responding Optimally

To determine the amount of evidence that should be assigned to any particular feature of a choice word, the system needs to know whether the feature is inactive or active and which sources might have caused activation. If the prime is a potential source of activation (i.e., if it is known that an activated feature has been primed), then the evidence assigned to this feature should be less because the prime, rather than the target, may have been the source of activation. We term this downward adjustment of evidence *discounting*.[5] To calculate the level of discounting in an optimal fashion, the system needs to know the activation probabilities α, β, and γ. These activation probabilities are not generally quantities that can be known to the decision system, because, among other things, they depend on details of the particular task. Thus, the evidence calculations need to be carried out with estimates of these quantities. The estimates the system uses to discount and, more generally, to evaluate evidence are termed α', β', and γ', respectively (these estimates turn out to be very close to the true values, but the difference between α and α' is a critical element of the model).

A Bayesian calculation yields the following odds (i.e., ratio of posterior likelihoods) for the target choice over the foil choice (assuming that each feature contributes an independent source of evidence; Huber et al., 2001):

$$\Phi\left(\frac{\mathrm{T}}{\mathrm{F}}\right) = \frac{\prod_{i=1}^{N} \frac{p(\mathrm{V}\{\mathrm{T}_i\} | \mathrm{T\ is\ target})}{p(\mathrm{V}\{\mathrm{T}_i\} | \mathrm{T\ is\ foil})}}{\prod_{j=1}^{N} \frac{p(\mathrm{V}\{\mathrm{F}_j\} | \mathrm{F\ is\ target})}{p(\mathrm{V}\{\mathrm{F}_j\} | \mathrm{F\ is\ foil})}}, \tag{1}$$

where T refers to the target word, F refers to the foil word, and $\mathrm{V}(\mathrm{T}_i)$ and $\mathrm{V}(\mathrm{F}_i)$ represent binary values denoting the state of activation of the i-th feature of the target and the foil, respectively. The products in the numerator and denominator of Equation 1 are the posterior likelihood that the target and the foil, respectively, were presented during the target flash, as determined from the N word features. When equal prior probabilities are assumed (as is appropriate for most experimental designs), a normative decision process will choose the word with the greater likelihood ratio (in the case of equal likelihood ratios, a random decision is made).

As Huber et al. (2001) pointed out, there are only three possible expressions for the product terms (i.e., the evidence provided by

[5] This is the same concept as *explaining away* in the study of Bayesian belief networks (i.e., the activation of the feature is "explained away" by the prime, and therefore the probability that the activation was due to the target is less; see, e.g., Wellman & Henrion, 1993).

Feature appeared in prime(s)	OFF (State of feature activation)	ON
NO	$\frac{(1-\gamma)(1-\beta)}{(1-\gamma)} = (1-\beta)$ less than 1.0	$\frac{1-(1-\gamma)(1-\beta)}{1-(1-\gamma)}$ greater than1.0
YES	$\frac{(1-\gamma)(1-\alpha')(1-\beta)}{(1-\gamma)(1-\alpha')} = (1-\beta)$ less than 1.0	$\frac{1-(1-\gamma)(1-\alpha')(1-\beta)}{1-(1-\gamma)(1-\alpha')}$ greater than, but closer to, 1.0

Figure 2. The evidence provided by features, contingent on the two states of activation and the two states of priming. The numerator in each fraction is conditional on the feature being shared with the target (and not the foil); the denominator is conditional on the feature being shared with the foil (and not the target). Features that are shared between the target and the foil do not provide discriminating information and are ignored in the decision process. β and γ refer to the probabilities of feature activation by the target presentation and visual noise, respectively. In theory, their estimates (β' and γ') should be used, but as discussed in the text, β and γ are assumed to be estimated correctly. α' denotes the estimated probability of feature activation by the prime presentation.

each feature) of Equation 1. These expressions depend on the state of activation of the feature and whether a prime is a potential source of activation. Figure 2 presents the expressions for feature evidence for the four combinations of feature activation and priming.

The numerators in the fractions of Figure 2 are the probability that a feature has its respective state of activation given the potential sources of activation, assuming the feature belongs to the target. The denominators are this same probability, assuming the feature does not belong to the target. For instance, consider the upper left-hand panel, in which the feature has not been primed and is inactive. In the numerator, the target flash and noise could have activated the feature, yet both of these sources must have failed to activate the feature, and so the probability is $(1 - \gamma)(1 - \beta)$. In the denominator, only noise could have activated the feature, and so the probability is $1 - \gamma$. The value of this fraction determines the amount of evidence contributed by each feature. The more this ratio differs from 1, the more evidence is contributed by the corresponding feature (the evidence is against the respective choice word when the ratio is less than 1 and for the respective choice word when it is larger than 1). As can be seen in Figure 2, activated features that are known to have appeared in the prime(s) are assigned a discounted level of evidence compared with activated features that did not appear in a prime. The critical difference between the upper right-hand fraction and the discounted fraction in the lower right-hand panel is the inclusion of the prime as a potential source of activation (i.e., the inclusion of the $1 - \alpha'$ terms). Therefore, α' is the primary determinant for the strength of discounting.

The qualitative predictions of ROUSE are not much affected by differences between β and its estimate β' or between γ and its estimate γ', so these estimates were set equal to the actual values (i.e., $\beta' := \beta$ and $\gamma' := \gamma$). However, model behavior critically depends on the amount and direction of the difference between the true probability of prime activation (α, i.e., the main determinant of source confusion) and its estimate (α', which determines discounting). For long prime durations, fitting the model revealed that $\alpha' > \alpha$, in which case primed features are discounted too heavily, resulting in a preference against primed choice words. For short prime durations, fitting the model revealed that $\alpha' < \alpha$, in which case primed features are not discounted enough, resulting in a preference for primed choice words.

Huber, Shiffrin, Quach, & Lyle (2002) provided evidence for another assumption of the model, that discounting of evidence should only apply to activated features known to be shared with a prime: The direction and degree of priming was strongly determined by the ability to identify the primes, as revealed by a prime-recognition test immediately following forced choice identification. When the primes are readily identifiable (as was likely the case in all experiments reported by Huber et al., 2001), it is assumed that it is known which features of the choice words were primed and, therefore, which features to discount.[6] However, at very brief prime durations, the prime might not be identifiable, and prime features may not be accessible to the discounting process. In this case, α' equals 0, and there is no discounting, such that even a small amount of source confusion could have a sizable effect. The evidence from the present experiments bears on this hypothesis.[7]

Experiment 1

We designed Experiment 1 to investigate (a) how variations in spatiotemporal separation between prime and target affect priming and (b) how near-threshold and above-threshold primes combine to produce priming effects. The former issue was addressed by crossing two levels of spatial separation (0 vs. 1 line) with two

[6] We draw a distinction between the system's knowledge of primed features and conscious experience of the prime. It is reasonable to assume that these concepts are correlated and, particularly, that once a prime can be identified, its features are accessible to the discounting process. However, we do not claim that conscious experience is necessary for discounting to take place or that discounting for primes that can be identified is necessarily excessive.

[7] Huber and O'Reilly (2003) reformulated ROUSE as a Bayesian network, demonstrating that a diminished certainty as to which features to discount can perform the same role as underestimation of the prime-activation probability. It is unclear which of these mathematically equivalent interpretations of underdiscounting is appropriate for very short prime durations.

levels of temporal separation (50 vs. 100 ms) between termination of the prime and onset of the target flash.[8] We investigated the latter issue with three conditions, using a *long-duration* prime (1,000 ms), a *short-duration* prime (50 ms), and a *long-then-short* prime (which was presented first for 950 ms and then again for 50 ms at a different spatial position).

It was unclear how two presentations of the same prime word would combine to affect source confusion and discounting. A simple combination rule might treat the amount of source confusion and discounting for the two back-to-back prime presentations as equivalent to that produced by a single presentation with the combined duration of both presentations (in this case, the results for the 1,000-ms and 950+50-ms prime-presentation conditions should be identical). In an alternative combination rule, the two prime presentations might combine independently, producing effects due to the summation of separate levels of source confusion and discounting (in this case, underdiscounting of the brief prime presentation should counteract the overdiscounting of the long prime presentation).

It is easiest to discriminate alternative hypotheses when experimental effects are largest. Huber (2005b; see also Huber & O'Reilly, 2003) found increased priming effects through the use of a single prime (presented simultaneously in two locations, above and below fixation) rather than two different primes. We therefore adopted this procedure.

Method

Participants. Sixty-nine undergraduate students at Indiana University Bloomington participated in exchange for introductory psychology course credit.

Materials. We used the same pool of 1,000 five-letter words used in previous repetition priming experiments by Huber, Shiffrin, Quach, and Lyle (2002) and Huber, Shiffrin, Lyle, and Quach (2002). The minimum written-language frequency of these words was 4 per million, as defined and measured by Kučera and Francis (1967). All words were presented in uppercase, using a fixed-width, 22-point font. The pattern mask for the target consisted of a row of six @ signs. Its height was adjusted (19 point, boldface) to correspond to the height of the letters, and its presentation was centered over the area where the target had been presented.

Equipment and display. All stimuli were displayed on 17-in. (43.18-cm) PC CRT monitors with a vertical refresh rate of 120 Hz and a screen resolution of 800 × 600 pixels. The display was synchronized to the vertical refresh using the ExpLib programming library (Cohen & Sautner, 2001). This provided display increments of 8.33 ms.

The stimuli were presented in white against a black background. Each participant sat in an enclosed booth with dim lighting. The distance of the monitor, the presentation positions, and the font size were chosen such that the target and the 50- and 1,000-ms primes encompassed less than 3° of vertical visual angle, and the 950-ms primes encompassed between 3° and 4° of vertical visual angle. The horizontal visual angle encompassed by the primes and the target was around 3°. Each prime presentation consisted of the same prime word presented simultaneously in two locations: symmetrically above and below fixation (see Figure 3), thus minimizing any tendency for eye movements to the prime.

Responses for the 2-AFC test were collected through a standard computer keyboard. Participants were asked to press the *Z* key or the *slash* key to choose the left or right alternative, respectively.

Procedure. Four factors were crossed within subject. Figure 3 illustrates the different conditions.

There were three priming conditions: neither primed (i.e., the prime was different from both choice words), target primed (i.e., the prime was identical to the target), and foil primed (i.e., the prime was identical to the foil). Words were randomly sampled without replacement such that each word only appeared on one trial in the experiment (nonrepetition of words across trials eliminated the possibility of contamination from long-term repetition priming).

There were six spatiotemporal conditions for priming: For each of the following three conditions, the two vertically aligned presentations of the (last) prime word were presented either with no separation (the bottom of the top presentation was touching, but not overlapping with, the top of the bottom presentation) or with one text-line separation (one row of mask characters separated the two vertically aligned prime presentations). In all cases, the most recent primes were followed immediately by a large mask of @ signs. The prime word was presented (a) for 1,000 ms, (b) for 50 ms, or (c) first with high vertical eccentricity for 950 ms and then again more centrally for an additional 50 ms, as in the 50-ms condition. Participants were instructed that in the 950+50-ms condition, the same prime word would appear in two separate back-to-back presentations (they were told that the prime word would move to a more central position for a second, brief presentation). Finally, the delay between prime offset and target onset (henceforth termed *delay*) was either 50 ms or 100 ms (see Footnote 8). During the delay, a mask was presented consisting of five rows of @ signs presented in vertical alignment with no space between the rows and no spaces between characters. Its presentation was centered over the display such that the entire area where primes could appear was covered.

Prior to the prime display, a fixation point was presented at the center of the screen for 500 ms. Following the mask that separated the primes from the target, the target was presented centrally (for an individually adjusted duration, as described below), followed by a central mask consisting of a row of @ signs. The mask duration was 500 ms minus the duration of the target for that trial (thereby keeping the interval from target onset to choice onset constant). The postmask was immediately followed by two choices, presented to the right and left, with target and foil positions randomly determined on each trial.

The first 64 trials of the experiment were used to adjust the time of the target presentation such that performance was at roughly 75% accuracy. Of the 64 calibration trials, half were neither primed, one quarter were target primed, and the remaining quarter were foil primed. The other variables were randomly assigned across trials.

Across participants, the average presentation time of the target was 50 ms, but as in previous studies (e.g., Huber, Shiffrin, Lyle, & Quach, 2002; Huber et al., 2001; Huber, Shiffrin, Quach, & Lyle, 2002), there were large individual differences. Target-flash times ranged from 25 ms to 108 ms across participants.

Following the block of 64 calibration trials, there were five blocks with 72 experimental trials each. In every block, each of the 36 conditions—3 (prime condition) × 3 (prime presentation) × 2 (separation) × 2 (delay)—was presented twice in a new random order.[9] Feedback was given after

[8] We conducted a pilot experiment that was identical to Experiment 1 with the exception that the two levels of temporal separation were 0 ms and 50 ms. However, the results of this experiment were difficult to interpret due to an unforeseen confound. In the 0-ms condition, the prime was masked by the target in the no-separation condition, but the prime was not masked in the separated condition because the separation was great enough to place the primes fully above and below the target. To avoid this confound, we did not include a condition with no temporal separation between prime and target in Experiment 1. In other words, in Experiment 1, the primes were always followed by pattern masks.

[9] Due to an oversight, one of the neither-primed trials in this experiment was replaced by a target- or foil-primed trial for 19 participants, and two neither-primed trials were replaced by target- and/or foil-primed trials for 6 participants, making the number of trials in the three prime conditions slightly uneven for 25 participants.

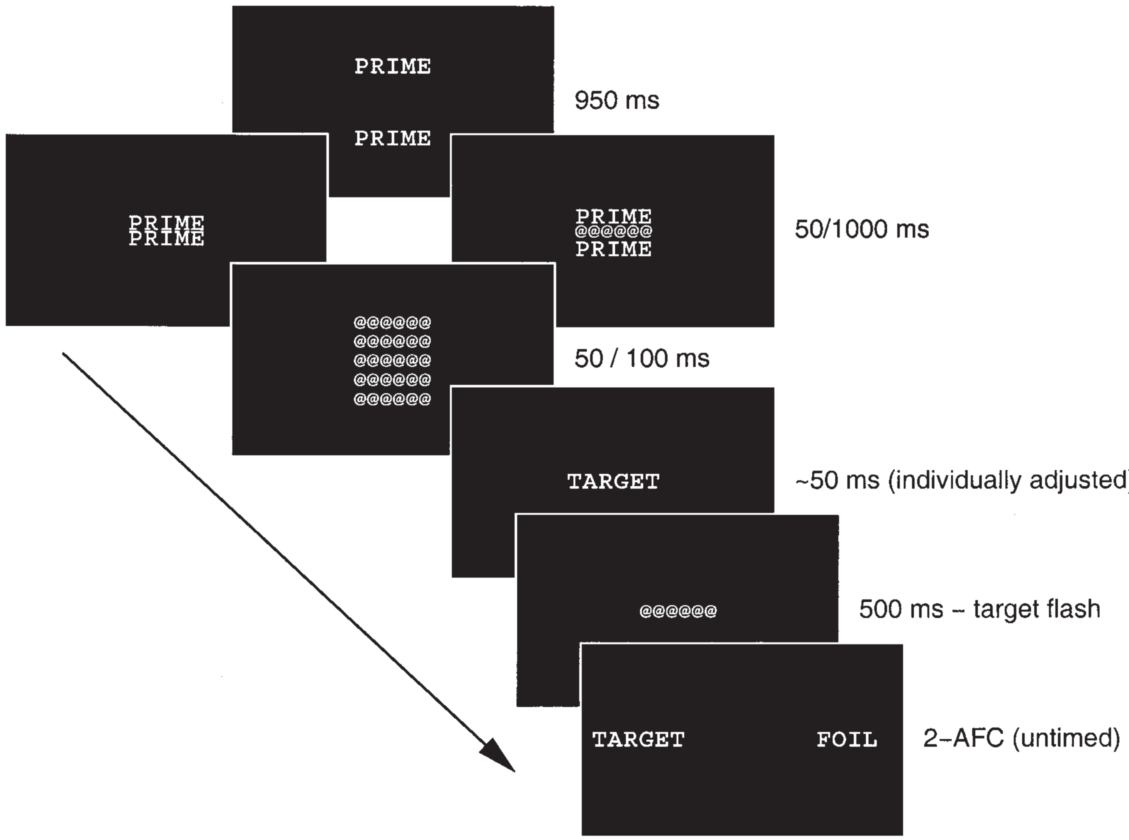

Figure 3. The sequence of events in Experiment 1. The 950-ms prime was presented only in the 950+50-ms conditions and was always followed by another 50-ms presentation. The first presentation of the prime was preceded by a 500-ms presentation of a fixation point in the center of the screen (not shown). The duration for each frame is indicated on the right. Only one of the two frames in the second row was shown on a given trial. The positions of the target and foil in the two-alternative forced choice (2-AFC) test were randomized.

every trial. A check mark and the word *correct* appeared in green when the answer was correct, and a cross mark (*X*) and the word *incorrect* were presented in red when the answer was incorrect. The feedback stayed on the screen for 700 ms and was immediately followed by the presentation of the fixation point for the next trial (unless the current trial was the last trial in a block).

After each block, participants received feedback providing the percentage of correct trials in the last block and the mean response time (this was the only time when feedback about response time was given, and the instructions emphasized accuracy rather than response speed). Participants were encouraged to take a short break between blocks and to resume the experiment only when they were ready to do so. The entire experiment took less than 40 min.

Results

To analyze the data, we carried out a 3 × 3 × 2 × 2 (Prime Condition × Prime Presentation × Separation × Delay) analysis of variance (ANOVA). All factors were within subject, and the dependent measure was accuracy. A significance level of .001 was used for all statistical tests. None of the statistical conclusions changed when the degrees of freedom of the *F* tests were adjusted to account for violation of the sphericity assumption (Geisser & Greenhouse, 1958). The bar heights in Figure 4 shows the results aggregated over the two levels of prime delay (none of the effects involving prime delay were significant). The dots in the figure give the ROUSE predictions, but discussion of the model predictions is deferred to the ROUSE Model Applied to Experiments 1–3 section. The numerical values for all conditions can be found in Appendix A. Before turning to the statistical analyses, we call attention to the fact that the results were not only reliable but quite large in magnitude, with probability of correct choice ranging from nearly .3 to over .9.

There were large effects for prime condition, $F(2, 136) = 104.62$, $MSE = 0.09$; prime presentation, $F(2, 136) = 76.18$, $MSE = 0.05$; and separation, $F(1, 68) = 118.11$, $MSE = 0.03$. It is clear from Figure 4 that the Prime Condition × Prime Presentation interaction, $F(4, 272) = 146.40$, $MSE = 0.04$, modulated these main effects: At the 50-ms prime presentation, performance was near ceiling when the target was primed, below chance when the foil was primed, and intermediate when neither the target nor the foil was primed. Presenting the prime for 1,000 ms radically changed this pattern of performance. The results reversed such that mean accuracy was lower for the target-primed condition than for the neither-primed condition, $t(275) = 4.835$, $SE = 0.013$. The results for the 950+50-ms conditions roughly represented an av-

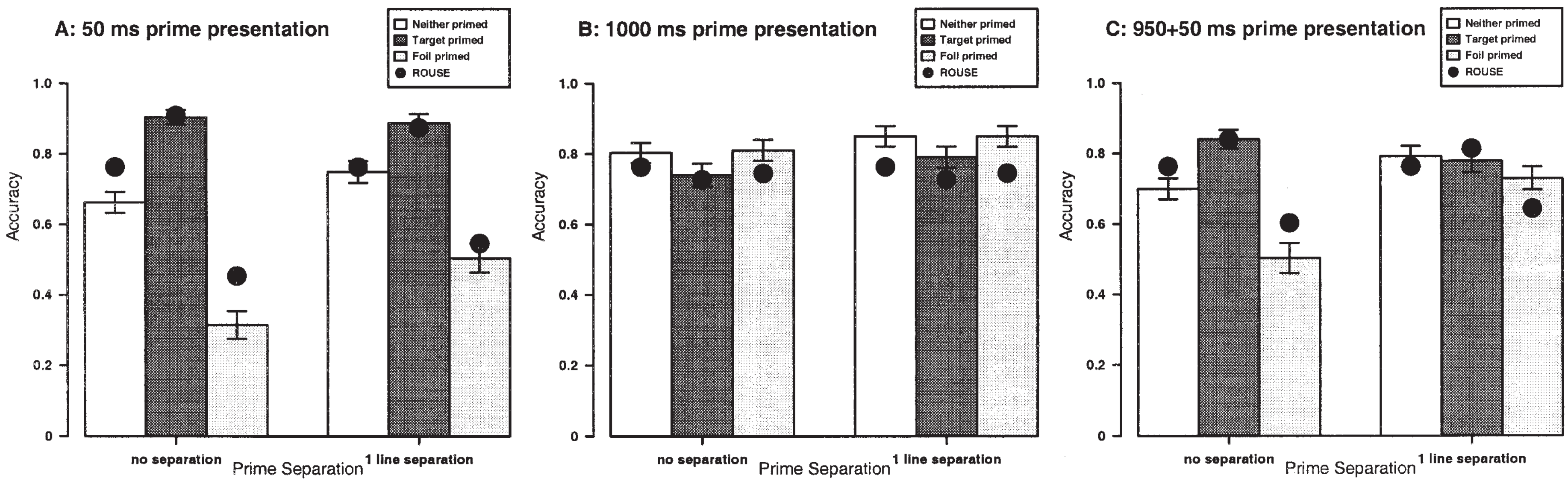

Figure 4. Results and corresponding ROUSE (responding optimally with unknown sources of evidence) predictions (represented by the dots) from Experiment 1: The bar heights show the mean proportions of correct choices (±95% confidence intervals [error bars]) for the three different prime-presentation conditions aggregated over the two levels of prime delay. The within-subject standard error, based on pooled error terms, is .019 (Loftus & Masson, 1994). The prime was always presented in two locations simultaneously, above and below, either with no separation or separated by one line of mask characters. Numerical values are given in Appendix A.

erage of the 50-ms and 1,000-ms conditions, subject to an interaction described next.

Finally, there were significant Prime Condition × Separation, $F(2, 136) = 63.43$, $MSE = 0.02$, and Prime Condition × Presentation × Separation, $F(4, 272) = 21.36$, $MSE = 0.02$, interactions. As can be seen from Figure 4, performance in the foil-primed condition was better when the primes were presented one line apart than when they were not, particularly in the 50-ms and 950+50-ms prime-presentation conditions. However, for the target-primed condition, performance with the 950+50-ms prime presentation was worse when the primes were separated compared with when they were not. For separated primes, mean accuracy for the target-primed conditions was not lower than that for the neither primed condition at the 950+50-ms prime presentation, $t(137) = 0.766$, $SE = 0.020$, $p = .22$.

Discussion

The pattern of results in the 50-ms and 1,000-ms prime-presentation conditions replicate those of Huber (2005b). For the 50-ms prime presentation, accuracy was nearly .6 higher when the target was primed than when the foil was primed (interpreted in ROUSE terms as underdiscounting of evidence from primed features); this contrasts with the 1,000-ms prime presentation, for which the difference between target priming and foil priming reversed (interpreted in ROUSE terms as overdiscounting of evidence from primed features). Note that the discounting process is graded in nature, and thus the absence of a priming reversal does not necessarily imply a lack of discounting but might, instead, imply that discounting was insufficient to fully or excessively counteract the effects of source confusion. However, a reversal in the direction of priming, such as occurred in the 1,000-ms prime presentation, provides unambiguous evidence that the discounting process is needed, because this result cannot be explained by varying levels of source confusion alone.

The important new results come from the 950+50-ms condition. The results in this condition were decidedly different than those in the 1,000-ms condition, instead appearing similar to the average of the 1,000-ms and 50-ms conditions. Although this pattern of results may on first sight be surprising, it can be explained with the simple assumption that the separate effects of a long and a short prime presentation combine independently (i.e., each separate presentation has an independent chance of producing source confusion and discounting for primed features). Presumably, the 50-ms condition produced some features that were confused with the target percept, and furthermore, these features were scarcely, if at all, discounted. Likewise, the 1,000-ms conditions produced features that were confused with the target percept, but these were counteracted by an even larger degree of discounting. According to an independent-sources interpretation of the 950+50-ms condition, the second presentation (i.e., 50 ms) provided additional source confusion with little or no additional discounting, resulting in a preference for the primed choice word despite the total 1,000-ms duration during which the prime was presented in some form.[10] This idea is implemented quantitatively in the ROUSE Model Applied to Experiments 1–3 section.

In their Experiment 1, using a 2-AFC perceptual identification task, Huber, Shiffrin, Quach, & Lyle (2002) contrasted a *long* priming condition, in which two prime words were presented in boldface for 2,500 ms, with a *long–switch* condition, in which two prime words were first presented regularly for 2,000 ms and then, following a position switch, presented for an additional 500 ms. Although the authors' *short* condition (i.e., a 500-ms presentation of two prime words) produced a preference for a primed alternative, this preference diminished in their long condition and was even further reduced in their long–switch condition. Superficially, their long–switch condition seems to have been similar to our long-then-short (950+50-ms) priming condition, so this last result is puzzling. However, because the second presentation of the primes was for 500 ms, as compared with 50 ms in the current situation, it may have been that the above-threshold second presentation resulted in the continual accrual of discounting. We assess the generality of this proposed theoretical account in Experiments 2 and 3.

Next, we consider the seemingly curious prime-separation effects. As seen in Figure 4, it appears that prime separation had a sizable effect for the 50-ms and 950+50-ms conditions but not for the 1,000-ms condition. More specifically, for both the 50-ms condition and the 950+50-ms condition, presenting the prime more centrally (i.e., not separated) resulted in a stronger preference for the primed alternative (i.e., magnification of the difference between the target-primed and foil-primed conditions). This result is consistent with the increased priming effect for central primes observed by Hochhaus and Marohn (1991, Experiment 4).

Surprisingly, we found no effect of prime separation in the 1,000-ms condition. We explain this with the post hoc assumption that source confusion for brief primes is more highly dependent on spatial proximity to the target than is source confusion for long primes. The idea is that long durations allow clear perception, largely independent of separation, and confusions are therefore based more on higher level cognitive and attentional factors. Thus, we modeled these results by assuming one level of source confusion for long prime presentations, regardless of where these presentations occurred, whereas we allowed separate levels of source confusion for short presentations, depending on the degree of visual separation (i.e., less source confusion for primes presented farther from fixation). Hochhaus and Marohn (1991, Experiment 4) presented a prime for 500 ms and still obtained a sizable separation effect, so the prime duration required to diminish this effect may be longer than 500 ms but shorter than 1,000 ms. However, if attentional (re)deployment to eccentric primes at longer prime durations is responsible for the diminished separation effect, then other variables (such as attentional load) are likely to modulate the influence of prime duration.

Finally, we consider explanations for the lack of effect for delay between prime offset and target onset (i.e., 50-ms vs. 100-ms mask following the primes). According to some views, the degree to which prime features and target features merge would depend on

[10] As graphed in Figure 4, the results for the separated 950-ms and 50-ms prime presentations seem to be similar to those for the 1,000-ms prime presentation. Indeed, the strongest support for the claim that separate presentations act as separate, independent sources comes from the non-separated prime presentations. Note, however, that even for the separated primes, performance is qualitatively different for the 1,000-ms presentation as compared with the 950+50-ms prime presentation: Only the 1,000-ms presentation resulted in a preference against a primed target.

the temporal interval separating them. However, the findings of Sanborn, Malmberg, and Shiffrin (2004) suggest that postmasking (in our case of both prime and target) tends to cause decisions to be based on higher level, more abstract features (e.g., letter identity) rather than on visual form. Such abstract features might survive short intervals like 50 ms and 100 ms relatively intact. Furthermore, it is plausible that the critical time at which confusions between target and prime play a role is when the two choices arrive. From this perspective, increasing the prime–target stimulus onset asynchrony from 50 ms to 100 ms would have little effect given the additional 500-ms delay before onset of the choice words. This view is in keeping with the neural version of ROUSE, proposed by Huber and O'Reilly (2003), which assumes that residual prime activation at the time of the choice word presentation is responsible for the source confusion.

Experiment 2

The results of the 950+50-ms conditions of Experiment 1 suggest that each temporally and spatially separate presentation of a prime can act as a separate source of source confusion and discounting, which then combine independently to produce an overall priming effect. This result is particularly important because it implies that each new presentation of a word is treated as novel source of information. During the early moments of a presentation, features may easily migrate and become confused with other presentations. With additional exposure time, features are still subject to source confusion, but identification of the presented item allows for corrective measures. This is a fundamentally different interpretation of feature binding than has been assumed by others (e.g., in feature integration theory; Treisman & Gelade, 1980; Treisman & Gormican, 1988). According to our data, as interpreted with the ROUSE model, features are essentially often "unbound," even following long exposures, but the decisional system can partially (or excessively) correct for the unwanted effects of unbound features by utilizing the expected degree of source confusion given the known (identified) sources.

If separate presentations provide separate and independent sources of source confusion and discounting, it should be possible to arrange matters such that the two choice words could receive different levels of source confusion and discounting through differential priming. In particular, it should be possible to maximize preference effects by causing underdiscounted source confusion for one choice word (i.e., by presenting it very briefly as a prime) and overdiscounted source confusion for the other choice word (i.e., by priming it for a long time).

Consider a new version of the 950+50-ms prime-presentation condition of Experiment 1: If, for example, a prime matching the foil is presented for the first 950 ms, this should lead to a preference against the foil (caused by overdiscounting of evidence), thus increasing performance. If a prime matching the target is then briefly presented near perceptual threshold but at a level sufficient to generate features and produce source confusions, those features might be given little or no discounting (because they often go unattended or unnoticed). This should also lead to a choice of target and, hence, should also increase performance. These two effects should summate, thereby producing an even larger enhancement of performance than that found in Experiment 1. The same reasoning, applied to the case in which the long prime matches the target and the short prime matches the foil, should result in unusually low performance as compared with the conditions in Experiment 1. Taken together, these predicted effects should amplify the already large preference effects found in Experiment 1.

Method

Participants. Forty-three undergraduate students at Indiana University Bloomington participated in exchange for introductory psychology course credit.

Materials and equipment. Materials and equipment for Experiment 2 were the same as in Experiment 1.

Procedure. Figure 5 illustrates the sequence of events in Experiment 2. Two factors were crossed within subject. The first factor was the nature of the first prime. This prime was presented for 950 ms with high vertical eccentricity (as was the first prime in the 950+50-ms condition of Experiment 1) and was identical to the target, the foil, or neither choice. The second factor was the nature of the second prime. This prime was centrally presented for 50 ms and also primed the target, the foil, or neither choice. Thus, there were a total of 9 (3×3) priming conditions. To maintain consistency with Experiment 1 and to maximize the priming effects, we presented each prime twice simultaneously (see Figure 5). In the condition in which both primes were different from the choice words (i.e., neither primed), the same unrelated word was presented during both prime presentations.[11] The second prime, presented twice simultaneously, was always presented in vertical alignment without an intervening line and with zero delay between the second prime and the target. Participants were made aware of the fact that the two prime presentations could be either the same word or different words.

The first 72 trials of the experiment were used to adjust the target-flash durations so as to place performance near 75% accuracy. The number of trials was equal for each of the nine priming conditions. Across participants, target presentation times ranged from 42 ms to 125 ms, with an average of 66 ms.

Following the calibration block of trials, there were five blocks of experimental trials, with 63 trials in each block. Within each experimental block, 7 trials of each condition were presented in a new random order. The rest of the procedure was identical to that described for Experiment 1.

Results

We analyzed the results of Experiment 2 with a 3×3 (First Prime $\times$ Second Prime) ANOVA. As in Experiment 1, the dependent measure was accuracy, and both factors were within subject. None of the statistical conclusions changed when the degrees of freedom of the F test were adjusted for violation of the sphericity assumption (Geisser & Greenhouse, 1958). The results are presented as bar heights in Figure 6, and the dots give ROUSE predictions. The numerical values corresponding to the bars and dots in the figure appear in Appendix B (discussion of the ROUSE predictions is deferred to the ROUSE Model Applied to Experiments 1–3 section).

The main effect of the first prime (950 ms) was significant, $F(2,$

[11] According to ROUSE, this should produce the same results as the presentation of two different, unrelated primes as long as it can be assumed that the unrelated primes share no features with either choice word. To the extent that an "unrelated" prime shares some features with the choice words randomly, performance is expected to be slightly lower for neither-primed conditions using two different primes.

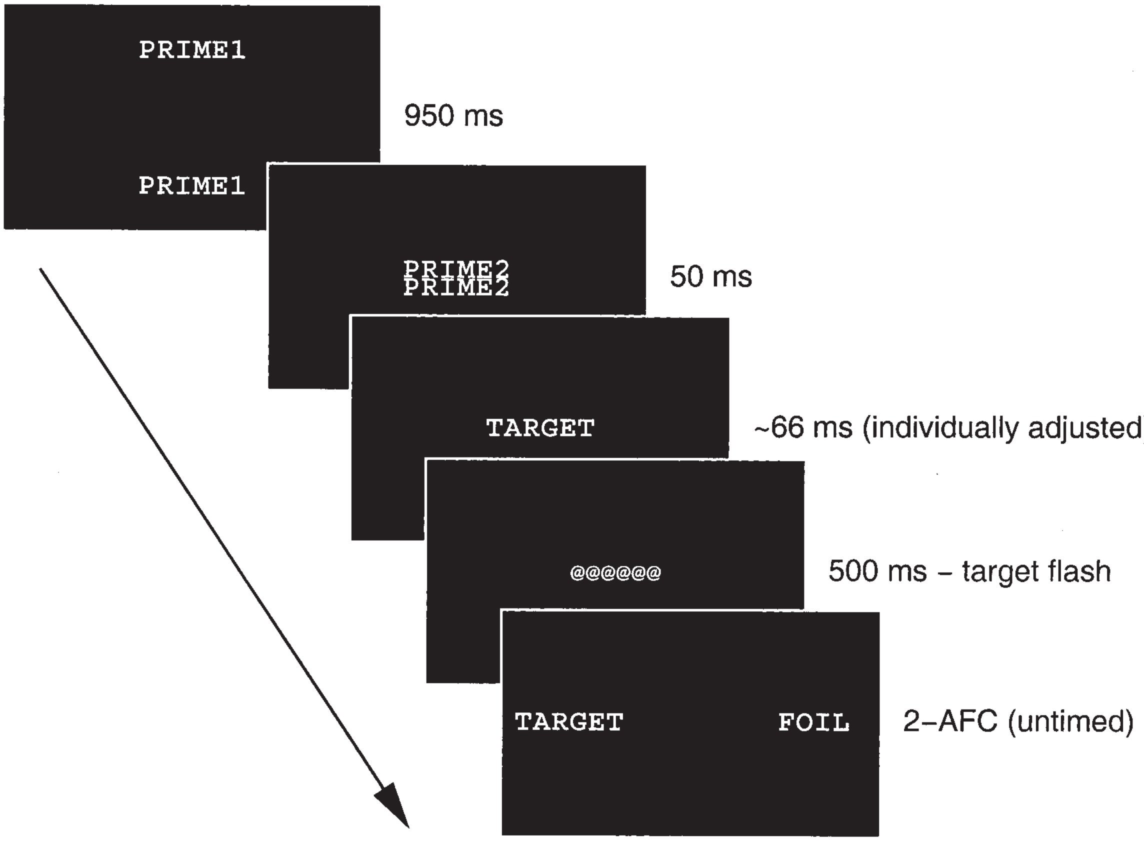

Figure 5. The sequence of events in Experiment 2. The first (950-ms) and second (50-ms) sequential prime presentations were independently varied (i.e., each could be the target, the foil, or neither, yielding 3 × 3 = 9 priming conditions). The presentation of the first prime was preceded by a 500-ms presentation of a fixation point in the center of the screen (not shown). The duration for each frame is indicated on the right. The positions of the target and foil in the two-alternative forced choice (2-AFC) test were randomized.

84) = 67.58, $MSE = 0.01$, $p < .001$. Performance was lowest when the target was presented as the first prime and higher when an unrelated word or the foil was presented. The main effect of the second prime (50 ms) also was significant, $F(2, 84) = 219.63$, $MSE = 0.03$, $p < .001$. In contrast to the first prime, performance was lowest when the second prime was the foil and considerably higher when it was the target. When the second prime was unrelated to both choice words, performance was intermediate. Finally, there was a significant First Prime × Second Prime interaction, $F(4, 168) = 33.77$, $MSE = 0.01$, $p < .001$. In conditions in which both the first and the second prime were expected to cause performance to increase (the first [long] prime matching the foil, the second [short] prime matching the target), the effects of the two primes amplified each other, resulting in performance near ceiling ($M = .91$). Similarly, in conditions in which both the first and the second prime were expected to cause performance to decrease (the first [long] prime matching the target, the second [short] prime matching the foil), the effects of both primes again amplified each other, resulting in performance near floor ($M = .26$). Most important, as predicted, the difference between the short conditions was larger than that between the long-then-short conditions, $t(42) = 6.466$, $SE = 0.036$, $p < .001$, and in turn, the difference between the conditions priming both choice words was larger than that between the short conditions, $t(42) = 6.073$, $SE = 0.032$, $p < .001$.

Discussion

The results of Experiment 2 generally conformed to expectations based on the assumptions that (a) the two prime presentations independently activate prime features and (b) matching features are underdiscounted in the decision process for a short prime presentation and overdiscounted in the decision process for a long prime presentation. In cases in which the target or the foil was presented as a first prime for 950 ms, followed by an unrelated word prime for 50 ms, performance was about equal (i.e., the long conditions, depicted in Bars 2 and 3 in Figure 6). Extrapolating from the results for 1,000-ms prime presentation in Experiment 1, we expected a small preference against the primed alternative. Although this trend did not reach significance, this does not necessarily suggest a lack of discounting. Lower levels of discounting than those seen in the long conditions in Experiment 1 may only attenuate or eliminate the positive priming effect and fall short of switching the preference against the primed word. The intervening unrelated prime may have attenuated the excessive

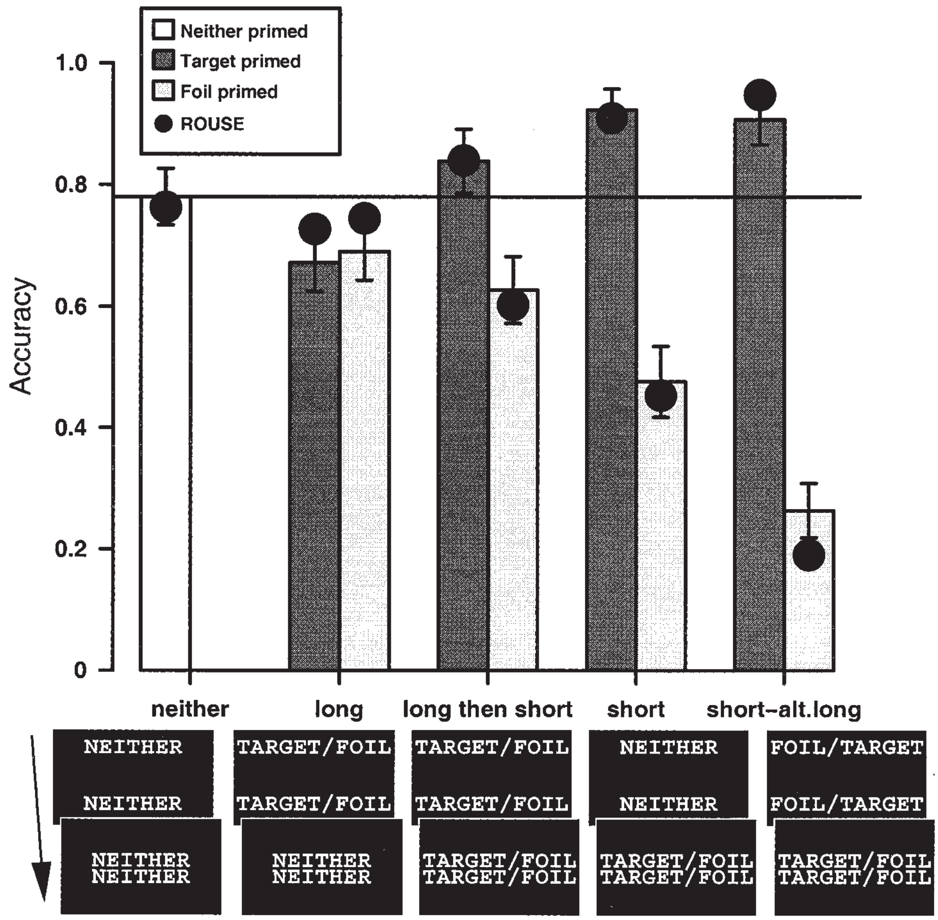

Figure 6. Results and corresponding ROUSE (responding optimally with unknown sources of evidence) predictions (represented by the dots) from Experiment 2: The bar heights show the mean proportions of correctly identified targets (±95% confidence intervals [error bars]) for the nine priming conditions. The within-subject standard error, based on a single-factor repeated measures analysis of variance, is .019 (Loftus & Masson, 1994). There was always a 950-ms (*long*) eccentric presentation of a prime word, immediately followed by a 50-ms (*short*) central presentation of the same or of a different prime word. In the *neither* condition, a word that was different from both choice words was presented during both the long and short durations. In the *long* conditions, the primed choice word was presented only during the long duration, and a word that was different from both choice words was presented for the short presentation. In the *long-then-short* conditions, the primed choice word was presented during both the long and short durations. In the *short* conditions, the primed choice word was presented only during the short duration (preceded by the presentation of a word different from both choice words during the long duration). Finally, the *short–alternative long* (short-alt.long) conditions primed both choice words differentially. For the last two bars, the labels *Target primed* and *Foil primed* refer to the briefly primed choice word, whereas the alternative choice word was presented for the long duration. The symbolized screens below the data illustrate the conditions: The top screens refer to the long prime presentation, and the bottom screens refer to the subsequent short prime presentation. For screens that include two words (i.e., *TARGET/FOIL* or *FOIL/TARGET*), the words before the slash apply to the conditions labeled as *Target primed*, and the words after the slash correspond to conditions labeled as *Foil primed*. Numerical values are given in Appendix B.

discounting seen in Experiment 1 just enough to eliminate the preference against the primed word. The significant difference between the short and long-then-short conditions, as well as that between the conditions priming both choice words and the short conditions, clearly suggests the operation of discounting.

For the long-then-short conditions (see Bars 4 and 5 in Figure 6), the primed choice word appeared during both prime presentations. As in Experiment 1, we assume that the second 50-ms presentation adds additional source confusion but little or no additional discounting, resulting in an increased preference for the primed alternative (compared with the long condition). For the short conditions (see Bars 6 and 7 in Figure 6), the primed choice word only appeared during the second 50-ms prime presentation, and as in Experiment 1, this resulted in a strong preference for the primed alternative; we assume that this effect arose due to source confusion combined with little or no discounting.

Finally, for the conditions priming both choice words (see the last two bars in Figure 6), one choice word was primed by the 50-ms presentation, and the other choice word was primed by the 950-ms presentation. We assume that brief presentation of a target prime produces underdiscounting of target features and that long presentation of a foil prime produces overdiscounting of foil features. These preference effects combine to produce nearly perfect accuracy. We assume the opposite occurs when the target is primed by a long prime and the foil is primed by a short prime, producing extremely low accuracy (in fact, accuracy greatly below the chance level of 50%). In other words, decisions were based more strongly on these priming effects than on valid perception of the briefly presented target.

Experiment 3

The results from Experiments 1 and 2 suggest that sequential prime presentations independently activate prime features, which are then discounted on the basis of prime saliency for each separate presentation. The effects of these separate presentations sum independently to determine the preference for each choice word. To ascertain whether prime saliency, rather than simple prime duration, is truly the underlying variable, in Experiment 3 we kept the prime duration constant while manipulating saliency through selective attention. According to the salience hypothesis, instructions to attend to one prime while ignoring the other should produce results analogous to those found for primes with different durations. We expected that instructions to attend to a prime would increase its saliency, just as does increasing its presentation duration, and that highly salient primes would result in greater discounting. Thus, features from attended primes should be overdiscounted, whereas features from ignored primes should be underdiscounted. Moreover, analogous to our findings in Experiment 2, we expected these preference effects to combine independently when both choice words were differentially primed.

Method

Participants. Forty-three undergraduate students at Indiana University Bloomington participated in exchange for introductory psychology course credit.

Materials and equipment. Materials and equipment for Experiment 3 were the same as in the previous two experiments.

Procedure. Figure 7 illustrates the sequence of events in Experiment 3. Each trial started with the presentation of two primes above and below fixation, one of which was shown in red, whereas the other was shown in blue. There was one line of separation between the two primes. Participants were instructed to ignore the blue prime and to make a judgment about the relative number of vowels and consonants in the red prime. Specifically, they were asked to press the *slash* key (labeled *RIGHT* +/− *1*) whenever the absolute difference between the number of vowels and consonants in the red prime equaled one and to press the *Z* key (labeled *LEFT* >/< *1*) otherwise. Participants were asked to count each vowel or consonant only once, even if it was repeated in the word. We chose this task because (a) it requires participants to pay attention to the orthography of the entire red prime, (b) it is sufficiently difficult to ensure that exposure to the primes is in the order of seconds, and (c) the pool of 1,000 words used in this experiment divides almost evenly between the two response categories (497 words required the *RIGHT* +/− *1* response).

Immediately after the classification of the red prime, the target word appeared briefly in the middle of the screen, followed by the usual pattern mask and 2-AFC test, just as in Experiments 1 and 2. Two different prime words were presented on every trial, and each prime could match the target, the foil, or neither choice word. Because there were two different primes on each trial, not all combinations could be realized. Specifically, a condition requiring the same prime to be attended and ignored (this would be analogous to the long-then-short condition of Experiment 2) was not possible. This led to seven priming conditions (see Figure 8): neither primed, attended prime matched target, attended prime matched foil, ignored prime matched target, ignored prime matched foil, ignored prime matched target and attended prime matched foil, and ignored prime matched foil and attended prime matched target.

The first 75 trials of the experiment were used to adjust the target-flash duration so as to place performance near 75% accuracy. For these calibration trials, all primes were different from the choice words (i.e., neither primed). Across participants, target presentation times after calibration ranged from 8 ms to 158 ms, with an average of 63 ms.

Following the calibration block of trials, there were five blocks of experimental trials, with 49 trials in each block. Within each experimental block, 7 trials of each condition were presented in a new random order. Feedback for both the consonant–vowel task and the perceptual identification task was given after every trial. Two check marks and the messages *You classified the word correctly!* and *You identified the word correctly!* appeared in green when the priming and the perceptual identification tasks were both solved accurately. If the response on either task was incorrect, the respective check mark was replaced by a red cross mark, and the respective message was changed to *You did NOT classify the word correctly!* or *You did NOT identify the word correctly!,* which also appeared in red. The feedback stayed on the screen for 1,800 ms and was immediately followed by the next trial (unless the current trial was the last trial in a block). After each block, participants received feedback on their percentages of correct classifications and identifications in the block. They were encouraged to take a short break and could resume the experiment when they were ready to do so. The entire experiment lasted approximately 45 min.

Results

Mean accuracy for the priming task on the red prime was .848 (SE = .009) across participants and conditions. Participants took, on average, 3,085 ms (SE = 62) to complete the priming task. The results for the perceptual identification task are presented as bar heights in Figure 8; the dots give the ROUSE predictions. The numerical values appear in Appendix C (the ROUSE predictions are discussed below in the ROUSE Model Applied to Experiments 1–3 section). Because not all theoretically possible combinations of the variables attended (neither/target/foil) and ignored (neither/target/foil) were realized in this experiment (due to the restriction that two different primes be presented on each trial), we tested specific contrasts for statistical significance. There was no significant difference between the attended target-primed condition and the attended foil-primed condition, $t(42) = 0.107$, $SE = 0.037$, $p = .46$. There was significantly higher accuracy for the ignored target-primed condition than for the ignored foil-primed condition, $t(42) = 4.752$, $SE = 0.135$, $p < .001$. A similar finding obtained in a comparison of the target-ignored–foil-attended condition and the foil-ignored–target-attended condition, $t(42) = 6.160$, $SE = 0.036$, $p < .001$. It is important to note that the difference between these conditions was larger in the latter case, $t(42) = 2.024$, $SE = 0.044$, $p < .05$.

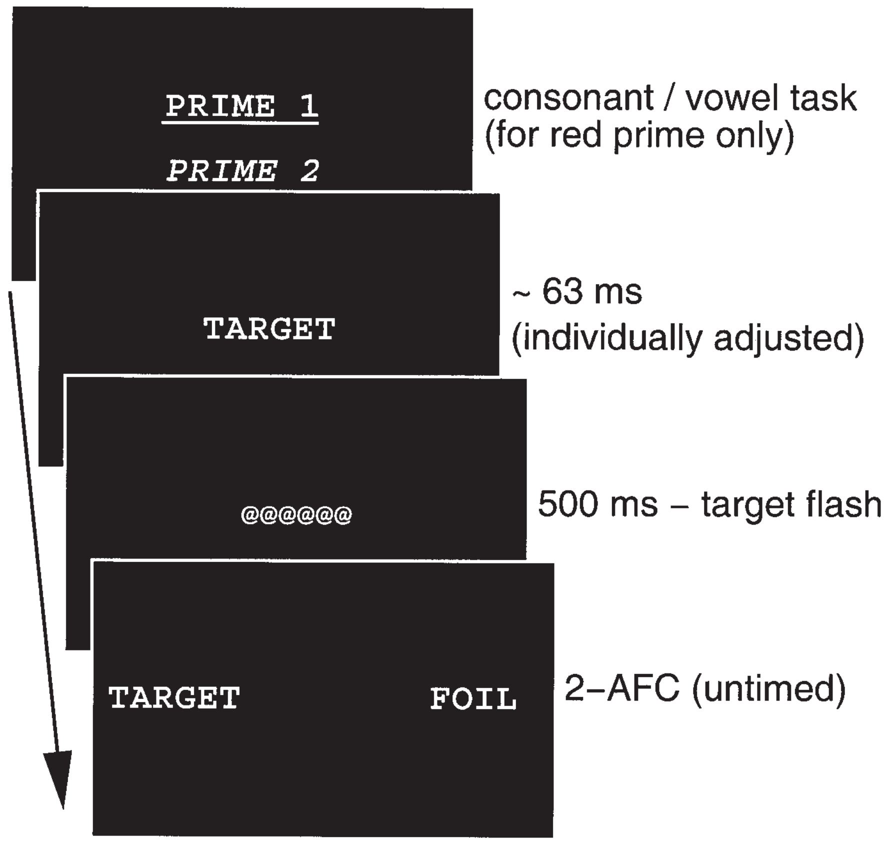

Figure 7. The sequence of events in Experiment 3. Primes were presented in the same font as targets, but one prime was shown in red (underlined in the figure), whereas the other prime was shown in blue (italicized in the figure; the up–down positions of the red and blue primes were randomly counterbalanced). The primes stayed onscreen for the duration of the consonant–vowel task (i.e., the selective-attention manipulation), which was in relation to the red prime (M = 3,085 ms). Once a response was given, the primes were immediately replaced by the target flash. Primes were repetitions of the target, the foil, or neither choice word. Two unique primes were presented on each trial, thus eliminating certain combinations (i.e., it was not possible for a single prime to be attended and ignored on the same trial). The durations for the other frames are indicated on the right. The positions of the target and foil in the two-alternative forced choice (2-AFC) test were randomized.

Discussion

Experiment 3 demonstrated that the effects of prime duration and the combination of multiple primes of different durations generalize to the case in which attention to the primes, rather than prime duration, is manipulated. Analogous to the results of Experiment 2, the results of Experiment 3 generally conform to expectations resulting from the assumptions that (a) both primes independently activate prime features and (b) matching features are underdiscounted in the decision process if the prime is ignored and overdiscounted in the decision process if the prime is attended. Hence, decisions in the 2-AFC test were biased in favor of ignored primes. This replication of the results from Experiment 2 is important, because many studies have used similar techniques to study selective attention (e.g., Sperling, Wurst, & Lu, 1993; see also Tipper, 2001).

The lack of preference against an attended prime does not provide evidence against the proposed discounting of attended prime features. Analogous to our argument regarding the lack of preference against the primed choice word in the long condition of Experiment 2, we believe that discounting of attended prime features in the present experiment was not strong enough to cause a preference reversal (as was seen in Experiment 1) but that, rather, discounting was merely sufficient to eliminate a positive priming effect. The claim of discounting for attended primes finds support in an examination of conditions in which one choice was primed by an attended prime whereas the other was primed by an ignored prime: If there were no effect of an attended prime, performance in these conditions (see the rightmost bars in Figure 8) should have been identical to that in the ignored conditions. Instead, priming the alternative choice by means of an attended prime magnified the preference effect, as would be the case if attended primes resulted in discounting. This pattern of results is highly similar to that of the short conditions and the conditions priming both choice words in Experiment 2 and thus supports the generality of our findings.

Previous studies using this paradigm have also manipulated attention to the primes, but these studies have confounded the attentional instructions with prime duration (Huber et al., 2001). Whereas Huber, Shiffrin, Quach, and Lyle (2002) established that

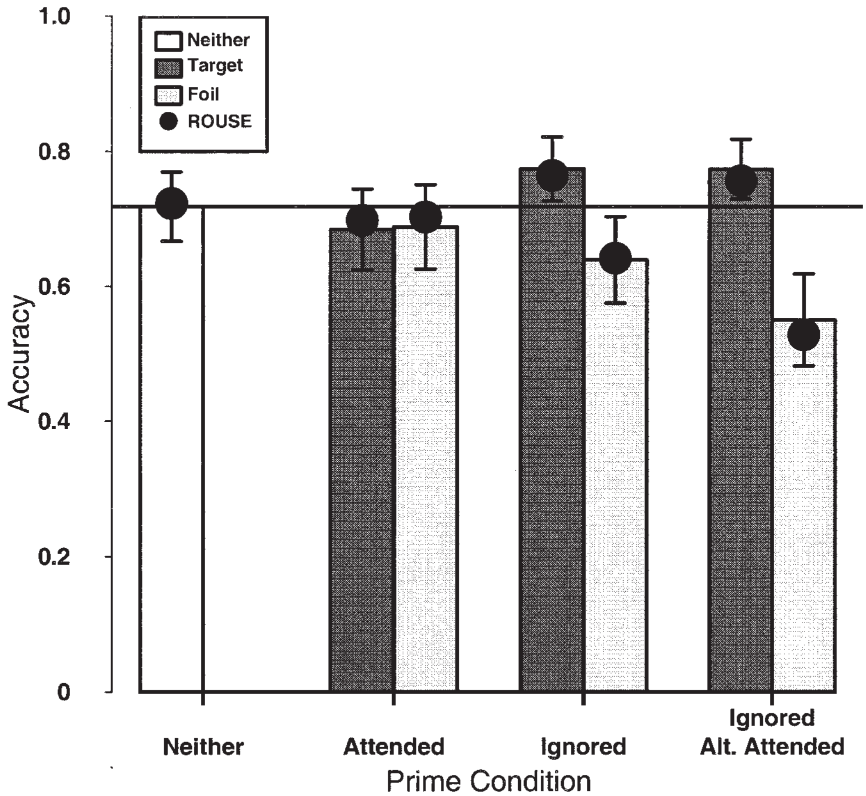

Figure 8. Results and corresponding ROUSE (responding optimally with unknown sources of evidence) predictions (represented by the dots) from Experiment 3: The bar heights show the mean proportions of correctly identified targets (±95% confidence intervals [error bars]) for the seven priming conditions. The within-subject standard error, based on a single-factor repeated measures analysis of variance, is .020 (Loftus & Masson, 1994). The color of the last two bars refers to the prime that was ignored while the other prime was attended to. Numerical values are given in Appendix C. Alt. = alternative.

prime duration alone can account for the differences between active and passive priming, the present experiment is the first strong demonstration that attentional instructions also influence discounting. Under the assumption that attention to a prime and explicit memory for that prime covary positively, our results are consistent with those of Huber, Shiffrin, Quach, & Lyle (2002), who found excessive discounting for remembered primes and underdiscounting for forgotten primes, even when prime duration was held constant.

Although our experimental design resembles the designs of many studies demonstrating negative priming (see Fox, 1995, for a review), we did not find an inhibited response to the ignored prime when it reappeared as the target—in fact, the opposite was the case. Following an application of ROUSE to the data of our experiments, we return to this apparent discrepancy in the General Discussion.

ROUSE Model Applied to Experiments 1–3

Although all of the experiments presented here are conceptually similar in that they demonstrate how primes of different saliency influence perceptual identification, both alone and in combination, the nature of the prime presentations in Experiments 1 and 2 is quite similar, but it is different from that in Experiment 3. We therefore chose to fit the ROUSE model simultaneously to the data of Experiments 1 and 2 and to estimate parameters for Experiment 3 separately.

ROUSE Applied to Experiments 1 and 2

The assumption in ROUSE that primed features are discounted only makes sense in situations in which the system has knowledge of the features that are in the prime(s); in previous studies with 2-AFC perceptual identification (e.g., Huber et al., 2001), primes were presented well above threshold, so such knowledge was presumably available and discounting could occur. However, in situations in which the prime is presented very briefly, the system may have very little or no knowledge of the prime features, reducing or eliminating the use of discounting. The 50-ms prime presentations in the present experiments are not very salient, so we assumed that the decisional system did not discount such primes at all (i.e., we set the estimated probability that features are activated by the prime, α', to 0).

We assumed that prime separation changed the probability that features were activated by the prime (i.e., α) only for short prime durations. For long prime durations, the same α was assumed for both levels of separation (see the *Discussion* section of Experiment 1). In addition, the same α was applied to both the 1,000-ms prime-presentation conditions and the first 950 ms of the 950+50-ms conditions, because these durations hardly differed, and the results of Experiment 1 suggested that vertical eccentricity of the prime display does not matter at such durations. As in previous applications of the ROUSE model (e.g., Huber et al., 2001; Huber, Shiffrin, Quach, & Lyle, 2002), β' and γ' were set to their true values, and γ was fixed at .02.

Thus, we fit the ROUSE model simultaneously to the data from Experiments 1 and 2 by estimating the following five parameters:

1. the probability that a feature is activated by a short (i.e., 50-ms) prime presentation when the prime display is central (α_{Scen}),

2. the probability that a feature is activated by a short (i.e., 50-ms) prime presentation when the prime display is separated (α_{Ssep}),

3. the probability that a feature is activated by a long (i.e., 950- or 1,000-ms) prime presentation (α_{L}),

4. the estimated probability that a feature is activated by a long (i.e., 950- or 1,000-ms) prime presentation (α'_{L}), and

5. the probability that a feature is activated by the target flash (β).

The probability of feature activation across the two prime presentations for the 950+50-ms conditions, $\alpha_{950+50\mathrm{cen}}$ (i.e., the 50-ms presentation is central) and $\alpha_{950+50\mathrm{sep}}$ (i.e., the 50-ms presentation is separated), were calculated by probabilistic summation of the appropriate α's for the first and the second prime presentation, assuming independence,[12] as specified in Equation 2,

$$\alpha_{950+50\mathrm{cen}} = \alpha_{\mathrm{L}} + \alpha_{\mathrm{Scen}} - (\alpha_{\mathrm{L}} \times \alpha_{\mathrm{Scen}}), \tag{2}$$

and Equation 3,

$$\alpha_{950+50\mathrm{sep}} = \alpha_{\mathrm{L}} + \alpha_{\mathrm{Ssep}} - (\alpha_{\mathrm{L}} \times \alpha_{\mathrm{Ssep}}). \tag{3}$$

Because we assumed that only long prime presentations (950 ms or 1,000 ms) resulted in any degree of discounting, words that were presented for a long duration were assigned the same level of discounting (i.e., the estimate of prime activation, α'_{L}) regardless of whether these words were also presented a second time for 50 ms.

In the original version of the ROUSE model (Huber et al., 2001), each trial was simulated probabilistically, but recently a computationally more efficient analytic method of calculating ROUSE predictions was developed by Huber (2005a). We used this analytic method for the current application.

The maximum likelihood chi-square error measure that was used by Huber et al. (2001; see also Correction to Huber et al. [2001], 2001), was used to obtain parameter estimates. For the fitting routine, a direct search algorithm, as proposed by Hooke and Jeeves (1961), was used (the actual implementation of the algorithm by Johnson, 1994, was based on Algol pseudocode published by Kaupe, 1963, with the improvements suggested by Bell & Pike, 1966, and Tomlin & Smith, 1969).

The model predictions for the experiments are presented graphically in the Figures 4 and 6 and numerically in Appendixes A and B. The corresponding parameters for these predictions were as follows: $\alpha_{\mathrm{Scen}} = .084$, $\alpha_{\mathrm{Ssep}} = .062$, $\alpha_{\mathrm{L}} = .071$, $\alpha'_{\mathrm{L}} = .082$, and $\beta = .050$. One has to be careful not to attribute too much significance to the absolute α values. Huber (2005a) explored the parameter variance–covariance matrix, finding that the difference between the actual and the estimated probabilities of feature activation by the prime is crucial for a good fit (i.e., $\alpha - \alpha'$ is critical), but the absolute value of these probabilities is less important (when the absolute value of α is shifted, reestimates of the other parameter values often produce almost equally good fits). It is clear from the figures that the pattern of results is captured by the ROUSE model, with the assumption of independent summation of effects for multiple prime sources.

In evaluating the goodness of the fit, it should be stressed that to keep the number of free parameters to a minimum, we have made strong assumptions that are likely to be only approximately true. For example, the probability that a feature is activated by the target flash (β) was assumed to be the same across experiments and across the different conditions in each experiment. Different levels of performance in the baseline neither-primed conditions suggest that this assumption does not precisely hold. Small changes in the rate of feature extraction from the target could be due to slight differences in the calibration procedures; different forward masking characteristics in the different conditions; and, perhaps, attentional effects, but we chose not to allow free parameters to capture such differences. Instead, the simple model we have described fully captures the qualitative pattern of the results.

ROUSE Applied to Experiment 3

Because prime duration was considerably above threshold in all conditions of Experiment 3, we had no reason to believe that any condition was immune to discounting. Therefore, fitting the model to the data of Experiment 3 involved estimating the following five parameters:

1. the probability that a feature is activated by an ignored prime (α_{I}),

2. the probability that a feature is activated by an attended prime (α_{A}),

3. the estimated probability that a feature is activated by an ignored prime (α'_{I}),

4. the estimated probability that a feature is activated by an attended prime (α'_{A}), and

5. the probability that a feature is activated by the target flash (β).

Parameter estimation was done as described for Experiments 1 and 2, yielding the following parameter estimates: $\alpha_{\mathrm{I}} = .064$,

[12] An alternative way to model the data would be to assume that only one prime presentation probabilistically dictates the level of source confusion and discounting on any given trial. We cannot distinguish this alternative account from the account proposed here with the data from our experiments, but note that fitting this alternative model would require an additional parameter to determine the probability that each prime is the critical prime on a given trial.

$\alpha_A = .064$, $\alpha'_I = .043$, $\alpha'_A = .073$, and $\beta = .040$. Again, it is important to stress that these exact α values are not crucial for a good fit: Changes in the absolute values of the actual prime activation probability (i.e., α) can be compensated by corresponding changes in the values for the estimated prime activation probability (i.e., α') over wide ranges of the parameter space. Although the best fitting values for α_I and α_A were equal, we obtained good fits when constraining them to differ by varying amounts, provided that the estimated prime activation probabilities were adjusted accordingly.

The model predictions are shown graphically in Figure 8 and numerically in Appendix C. It is clear that the fit of the model to the data is almost perfect, capturing the data both quantitatively and qualitatively. However, the ratio of free parameters to data points is much larger in the present fit than for Experiments 1 and 2.

Alternative Models

Huber and O'Reilly (2003) developed a neural network model for short-term priming called *nROUSE* (a neural mechanism for responding optimally with unknown sources of evidence). This model can account for some results that ROUSE cannot explain, and it can be adapted to handle response times and accuracy results simultaneously. However, it is a more powerful model with more mechanisms, and its increased explanatory power is not needed for the experiments presented in this article.

As an alternative to ROUSE, Ratcliff and McKoon (2001) developed a multinomial model that has been successfully applied to similar short-term priming data. This model is similar to ROUSE at an abstract level in that it also contains the opposing mechanisms of source confusion and discounting. However, the feature representation and probabilistic activation in ROUSE allow for a priori predictions regarding similarity relations and duration manipulations and their effects on priming, whereas the multinomial model has had to be adjusted appropriately for each new data set collected (see Huber, Shiffrin, Lyle, & Quach, 2002, for a related discussion). In any case, the experiments presented in the current situation were not specifically designed to differentiate between the multinomial model and ROUSE.

General Discussion

The study of short-term repetition priming using forced choice perceptual identification has highlighted the importance of modeling decision processes in addition to perceptual processes. The experiments presented here and previous research (e.g., Huber, Shiffrin, Lyle, & Quach, 2002; Huber et al., 2001; Huber, Shiffrin, Quach, & Lyle, 2002; Ratcliff & McKoon, 2001) provide strong evidence that two offsetting components—one largely perceptual or memorial (source confusion) and one largely decisional (discounting; see Footnote 1)—are needed to fully explicate priming phenomena. To date, only a few theories have explicitly included decisional aspects of the priming task, (Huber & O'Reilly, 2003; Huber et al., 2001; Ratcliff & McKoon, 2001), but in general such effects are likely to exist in many situations, particularly when the identity of the prime is known.

Source confusion refers to information arising from prime presentations (and from visual noise) that is confused with information arising from the target presentation. Such confusion produces a tendency to choose the word that is similar to the prime(s). Discounting of information known to have been in primes is used to counteract this preference. The degree of discounting must be estimated for the system to calculate evidence in favor of the two choices, and the estimated values may not match the actual degree of source confusion. In general, brief or ignored presentations result in too little discounting; in fact, primes near or below the threshold of explicit identification produce some degree of source confusion but may produce no discounting. In this case, underdiscounting produces a preference to choose the primed alternative. Long-duration or highly attended prime presentations can produce overestimation of source confusion and, thus, overdiscounting. This produces a tendency to choose the unprimed alternative.

Prior to the current set of experiments, ROUSE did not specify the nature of sequential prime presentations at different locations or the effect of differential attention to primes. In advance of these experiments, it was not clear how such factors would affect either the source confusion or the discounting mechanisms, nor how these factors would combine across presentations. In particular, our results strongly constrain the manner in which primes can reinforce or compete with each other when primes are presented for different durations and at different locations or are attended differentially.

In Experiment 1, we presented the same prime word at two different points in time, each in a different visual location. The first presentation was long (950 ms) and the second very brief (50 ms). In other conditions, we presented only a single prime presentation for a long duration (1,000 ms) or a very brief one (50 ms). Very brief presentations resulted in source confusion but little or no discounting (cf. Paap, Chun, & Vonnahme, 1999, Experiment 1, who also found large preference effects for brief primes), whereas long presentations resulted in source confusion as well as excessive discounting. The condition that presented the same prime for both a long duration and then a brief duration was well explained using the ROUSE model with the simple assumption that the source confusion and discounting from each presentation sum together independently to determine overall performance levels. This experiment also demonstrated that brief delays between prime and target do not significantly reduce source confusion (presumably because high-level features survive short delays and/or the slight delays are inconsequential in comparison with the much longer delays until the choices appear) and that presenting primes in visual locations more distant from the fixation point can reduce source confusion for brief, but not long, prime presentations (presumably because longer presentations allow attention to be [re]deployed to the primes).

In Experiment 2, we further tested the findings from Experiment 1 by putting the different levels of discounting associated with brief and long primes in competition or cooperation with one another. This was done by priming each choice separately, with one receiving a long prime and the other a brief prime. In this case, the results imply that different levels of prime-induced preference can exist for each choice word and that these combine to determine overall performance. Briefly priming one choice word (i.e., inducing a preference for that choice word due to source confusion with too little or no discounting) and a long prime presentation of the other choice word (i.e., inducing a preference against that choice word due to excessive discounting) combined to produce very large preference effects.

The results of Experiments 1 and 2 can be interpreted in terms of salience, with longer presentations resulting in highly salient primes, which were therefore more strongly discounted in the decision process. To test this interpretation, we manipulated attention as another variable that should affect prime salience in Experiment 3. By including a prime identification task, we induced selective attention to one of two primes while holding prime duration constant across both primes. As with the duration manipulations of Experiments 1 and 2, we observed underdiscounting for ignored primes and excessive discounting for attended primes. As in Experiment 2, these effects combined across choices, increasing preference effects when both choice words were differentially primed.

Paap et al. (2000, Experiment 3) compared repetition priming for words of low and high frequency. In their experiment, a related (e.g., *junk*) or neutral (e.g., *xxxx*) prime presented for 2,000 ms was followed by a brief target (e.g., *JUNK*), and participants had to identify a particular letter of the target in a subsequent 2-AFC test (for this example, the choice words would be *JUNK* and *BUNK*). Paap et al. (2000) found an increase in performance for trials using related primes compared with trials using neutral primes for low-frequency targets only. For high-frequency targets, performance for trials using related primes was actually slightly worse than that for trials using neutral primes (although this difference was not statistically significant). Referring to this interaction with word frequency, Paap et al. (2000) concluded that "because no model predicts that an identity prime should interfere with target processing, it appears reasonable to treat the difference as noise" (p. 1697). With ROUSE, there is now a model that predicts that identity primes can elicit performance deficits under certain circumstances. If it can be assumed that high-frequency primes are more salient than low-frequency primes (all other things being equal), then this aspect of Paap et al.'s (2000) data can be readily explained within the ROUSE framework by positing that a higher level of discounting for the high-frequency primes eliminates the positive priming effect found for low-frequency primes. The fact that there was a preference for the primed low-frequency target and no statistically significant preference against the primed high-frequency target, despite the prolonged prime duration, may have been due to the high similarity of the choice words. As demonstrated by Huber et al. (2001), ROUSE predicts that discounting will be less effective when fewer features discriminate between the choice words. Furthermore, it is important to note that in Paap et al.'s (2000) Experiment 3, the delay between prime and target was relatively long—the prime and target were separated by a warning signal (*LOOK*) and a premask, each of which lasted 1,000 ms—and thus it is unclear to what extent a theory of short-term priming should even apply. At this point, our interpretation of Paap et al.'s (2000) result is highly speculative. Future research using a paradigm capable of indexing decisional aspects of the task will be needed to ascertain whether word frequency is yet another variable affecting salience, and therefore discounting, similar to our manipulations of prime duration and attention.

Repetition Blindness

Repetition priming deficits similar to the ones observed in some conditions in the present experiments are routinely found in rapid serial visual presentation (RSVP) identification tasks and are known as *repetition blindness* (e.g., Kanwisher, 1987; Park & Kanwisher, 1994). Huber et al. (2001) discussed how ROUSE could be adapted to an RSVP identification task to account for these deficits in terms of overdiscounting of evidence from primed features.

The term repetition blindness, however, is not only used for RSVP identification tasks but has also been applied to repetition priming deficits in the perceptual identification of targets in single-prime–target trials (e.g., Hochhaus & Johnston, 1996; Hochhaus & Marohn, 1991). The ROUSE model is directly applicable to these results and predicts repetition blindness whenever the prime is salient enough to cause sufficient overdiscounting of evidence from primed features.

Negative Priming

The results of the present experiments, as well as the predictions of the ROUSE theory, seem to be at odds with certain negative priming (NP) findings. In a typical NP experiment, a target and distractor are presented simultaneously, and the task for the participant is to make a response to the target while ignoring the distractor (which is distinguished from the target by some cue, such as color). NP refers to the finding that the response latency for a target in trial n is increased when that target served as distractor in trial $n - 1$ (a decrease in response latency—i.e., a positive priming effect—is usually found if the target on trial n also served as target on trial $n - 1$; for a review, see Fox, 1995).

If the target and distractor of trial $n - 1$ serve as primes for trial n, ROUSE predicts sizable discounting for features from the attended prime (i.e., the target on trial $n - 1$) and relatively little discounting for features from the ignored prime (i.e., the distractor on trial $n - 1$). This should cause a priming benefit for previous distractors and a priming deficit for previous targets, which is exactly the opposite of the NP phenomenon.

When evaluating the discrepancies between our results and those of studies finding NP, it is crucial to note several important differences. First of all, NP manifests primarily in response times. Almost all studies investigating NP use response time as the main dependent variable and are designed such that accuracy is nearly perfect. The only study we are aware of that found NP in accuracy data in which accuracy of the task was not close to ceiling is that of Neill and Terry (1995), although their task was quite different from our perceptual identification task. Santee and Egeth (1982) suggested that accuracy measures different perceptual processes for short and long target exposures and, furthermore, that accuracy for short target exposures (such as those used in our 2-AFC perceptual identification paradigm) measures processes distinct from those measured by response time. Second, NP is not always reliably obtained, and when it is obtained, the effect is usually small (around 20 ms). There is continued dispute regarding the necessary and sufficient conditions for obtaining the NP effect (Fox, 1995).

In several unpublished studies, we have tried to mimic conditions under which NP is usually obtained but to do so within our 2-AFC perceptual identification paradigm. In none of these studies did we find NP. At this point, it is unclear how best to reconcile the findings of NP (using response time paradigms) and the seemingly opposite result in our 2-AFC paradigm, and we leave such reconciliation to future research.

Conclusions

We started this article with discussion of a puzzle: How can one make sense of the finding that primes whose presence is hard to ascertain produce as much or more priming than primes well above threshold? We have explicated this puzzle by using a model in which both the magnitude and the direction of priming is a function of the difference between activation of confusable prime features and estimates of that activation. Of course, we have gone much further. Using the assumption that primes near threshold produce source confusion but little or no discounting, and that above threshold primes produce both source confusion and discounting (sometimes more discounting than confusion), we have been able to account for a host of not very intuitive findings, including the way that multiple primes of different saliency combine. It is important to note that ROUSE is not limited to repetition word priming and, indeed, has been successfully applied to experiments using associative, orthographic, and phonological priming (Huber et al., 2001; Huber, Shiffrin, Lyle, & Quach, 2002) as well as to experiments using faces instead of words (Huber, 2005b). The results of these studies suggest that source confusion and discounting are general mechanisms and are not limited to a specific task or stimulus class.

These results significantly advance understanding of short-term priming phenomena specifically and of information processing more generally. For instance, phenomena similar to those of perceptual source confusion and discounting have been observed in the literature on social cognition, in which they are known as behavioral assimilation and contrast (e.g., Dijksterhuis et al., 1998). These results and others suggest that compensatory mechanisms, such as discounting, play a critical role in directing behavior in a wide variety of tasks, and for a wide variety of levels of representation, ranging from low-level perception to high-level conception. From this wider analysis, compensation does not exclusively result from explicit strategic control but, rather, is viewed as a feature of human cognitive architecture to accommodate basic limitations, such as source confusion, in the attempt to guide behavior in an optimal, or near optimal, manner.

References

Ashby, F. G., Prinzmetal, W., Ivry, R., & Maddox, W. T. (1996). A formal theory of feature binding in object perception. *Psychological Review, 103,* 165–192.

Bell, M., & Pike, M. C. (1966). Remark on algorithm 178 [E4] direct search. *Communications of the Association for Computing Machinery, 9,* 684–685.

Cohen, A. L., & Sautner, M. (2001). ExpLib (Version 1.0.1 [beta]) [Computer programming library for Visual C++]. Retrieved from http://people.umass.edu/alc/explib/index.htm

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

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108,* 204–256.

Correction to Huber et al. (2001). (2001). *Psychological Review, 108,* 652.

Dijksterhuis, A., Spears, R., Postmes, T., Stapel, D., Koomen, W., van Knippenberg, A., & Scheepers, D. (1998). Seeing one thing and doing another: Contrast effects in automatic behavior. *Journal of Personality and Social Psychology, 75,* 862–871.

Dill, M., & Edelman, S. (2001). Imperfect invariance to object translation in the discrimination of complex shapes. *Perception, 30,* 707–724.

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

Forster, K. I., & Veres, C. (1998). The prime lexicality effect: Form-priming as a function of prime awareness, lexical status, and discrimination difficulty. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 498–514.

Fox, E. (1995). Negative priming from ignored distractors in visual selection: A review. *Psychonomic Bulletin & Review, 2,* 145–173.

Geisser, S., & Greenhouse, S. W. (1958). An extension of Box's results on the use of the *F* distribution in multivariate analysis. *Annals of Mathematical Statistics, 29,* 885–891.

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

Hochhaus, L., & Johnston, J. C. (1996). Perceptual repetition blindness effects. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 355–366.

Hochhaus, L., & Marohn, K. M. (1991). Repetition blindness depends on perceptual capture and token individuation failure. *Journal of Experimental Psychology: Human Perception and Performance, 17,* 422–432.

Hooke, R., & Jeeves, T. A. (1961). "Direct search" solution of numerical and statistical problems. *Journal of the Association for Computing Machinery, 8,* 212–229.

Huber, D. E. (2005a). *Computer simulations of the ROUSE model: An analytic method and a generally applicable technique for producing parameter confidence intervals.* Manuscript submitted for publication.

Huber, D. E. (2005b). *The time course of perceptual processing for words and faces.* Manuscript in preparation.

Huber, D. E., & O'Reilly, R. C. (2003). Persistence and accommodation in short-term priming and other perceptual paradigms: Temporal segregation through synaptic depression. *Cognitive Science, 27,* 403–430.

Huber, D. E., Shiffrin, R. M., Lyle, K. B., & Quach, R. (2002). Mechanisms of source confusion and discounting in short-term priming 2: Effects of prime similarity and target duration. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 1120–1136.

Huber, D. E., Shiffrin, R. M., Lyle, K. B., & Ruys, K. I. (2001). Perception and preference in short-term word priming. *Psychological Review, 108,* 149–182.

Huber, D. E., Shiffrin, R. M., Quach, R., & Lyle, K. B. (2002). Mechanisms of source confusion and discounting in short-term priming: 1. Effects of prime duration and prime recognition. *Memory & Cognition, 30,* 745–757.

Humphreys, G. W., Besner, D., & Quinlan, P. T. (1988). Event perception and the word repetition effect. *Journal of Experimental Psychology: General, 117,* 51–67.

Johnson, M. G. (1994). Nonlinear optimization using the algorithm of Hooke and Jeeves [C source code]. Retrieved from http://www.netlib.org/opt/hooke.c

Kanwisher, N. (1987). Repetition blindness: Type recognition without token individuation. *Cognition, 27,* 117–143.

Kaupe, A. F., Jr. (1963). Algorithm 178: Direct search. *Communications of the Association for Computing Machinery, 6,* 313–314.

Kučera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.

Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review, 1,* 476–490.

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide.* Cambridge, England: Cambridge University Press.

Marcel, A. J. (1983). Conscious and unconscious perception: Experiments on visual masking and word recognition. *Cognitive Psychology, 15,* 197–237.

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

Merikle, P. M., & Daneman, M. (1998). Psychological investigations of unconscious perception. *Journal of Consciousness Studies, 5,* 5–18.

Merikle, P. M., Smilek, D., & Eastwood, J. D. (2001). Perception without awareness: Perspectives from cognitive psychology. *Cognition, 79,* 115–134.

Meyer, D. E., & Schvaneveldt, R. W. (1976, April 2). Meaning, memory structure, and mental processes. *Science, 192,* 27–33.

Neill, W. T., & Terry, K. M. (1995). Negative priming without reaction time: Effects on identification of masked letters. *Psychonomic Bulletin & Review, 2,* 121–123.

Paap, K. R., Chun, E., & Vonnahme, P. (1999). Discrete threshold versus continuous strength models of perceptual recognition. *Canadian Journal of Experimental Psychology, 53,* 277–293.

Paap, K. R., Johansen, L. S., Chun, E., & Vonnahme, P. (2000). Neighborhood frequency does affect performance in the Reicher task: Encoding or decision? *Journal of Experimental Psychology: Human Perception and Performance, 26,* 1691–1720.

Park, J., & Kanwisher, N. (1994). Determinants of repetition blindness. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 500–519.

Pecher, D., Zeelenberg, R., & Raaijmakers, J. G. W. (1998). Does pizza prime coin? Perceptual priming in lexical decision and pronunciation. *Journal of Memory and Language, 38,* 401–418.

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

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

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

Ratcliff, R., & McKoon, G. (2001). A multinomial model for short-term priming in word identification. *Psychological Review, 108,* 835–846.

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.

Sanborn, A. N., Malmberg, K. J., & Shiffrin, R. M. (2004). High-level effects of masking on perceptual identification. *Vision Research, 44,* 1427–1436.

Santee, J. L., & Egeth, H. E. (1982). Do reaction time and accuracy measure the same aspects of letter recognition? *Journal of Experimental Psychology: Human Perception and Performance, 8,* 489–501.

Schooler, L. J., Shiffrin, R. M., & Raaijmakers, J. G. W. (2001). A Bayesian model for implicit effects in perceptual identification. *Psychological Review, 108,* 257–272.

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

Shiffrin, R. M., & Steyvers, M. (1998). The effectiveness of retrieval from memory. In M. Oaksford & N. Chater (Eds.), *Rational models of cognition* (pp. 73–99). Oxford, England: Oxford University Press.

Sperling, G., Wurst, S. A., & Lu, Z.-L. (1993). Using repetition detection to define and localize the processes of selective attention. In D. E. Meyer & S. Kornblum (Eds.), *Attention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience* (pp. 265–298). Cambridge, MA: MIT Press.

Stankiewicz, B. J., & Hummel, J. E. (2002). Automatic priming for translation- and scale-invariant representations of object shape. *Visual Cognition, 9,* 719–739.

Tipper, S. P. (2001). Does negative priming reflect inhibitory mechanisms? A review and integration of conflicting views. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 54*(A), 321–343.

Tomlin, F. K., & Smith, L. B. (1969). Remark on algorithm 178 [E4]: Direct search. *Communications of the Association for Computing Machinery, 12,* 637–638.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12,* 97–136.

Treisman, A. M., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review, 95,* 15–48.

Wellman, M. P., & Henrion, M. (1993). Explaining "explaining away." *IEEE Transactions on Pattern Analysis and Machine Intelligence, 15,* 287–307.

Appendix A

Mean Accuracy (With Standard Errors in Parentheses) and Corresponding ROUSE Predictions for Experiment 1

	50-ms prime			1,000-ms prime			950+50-ms prime		
Separation and delay	Neither primed	Target primed	Foil primed	Neither primed	Target primed	Foil primed	Neither primed	Target primed	Foil primed
None									
50 ms	.652 (.021)	.913 (.015)	.312 (.027)	.789 (.021)	.728 (.025)	.797 (.021)	.688 (.022)	.835 (.020)	.515 (.029)
100 ms	.672 (.022)	.896 (.015)	.319 (.029)	.816 (.020)	.753 (.021)	.823 (.021)	.713 (.020)	.845 (.019)	.490 (.032)
ROUSE	.763	.909	.454	.763	.727	.745	.763	.840	.603
1 text line									
50 ms	.771 (.023)	.900 (.017)	.531 (.030)	.828 (.023)	.770 (.021)	.845 (.021)	.807 (.020)	.783 (.023)	.753 (.021)
100 ms	.728 (.021)	.877 (.019)	.478 (.028)	.872 (.018)	.811 (.021)	.854 (.021)	.777 (.021)	.771 (.022)	.706 (.024)
ROUSE	.763	.875	.546	.763	.727	.745	.763	.814	.643

Note. Within-subject standard error (based on pooled error terms): .019 (Loftus & Masson, 1994). ROUSE = responding optimally with unknown sources of evidence.

Appendix B

Mean Accuracy (With Standard Errors in Parentheses) and Corresponding ROUSE Predictions for Experiment 2

	First prime (950 ms)/second prime (50 ms)								
Accuracy	N/N	N/T	N/F	T/N	T/T	T/F	F/N	F/T	F/F
Observed	.780 (.023)	.923 (.017)	.476 (.029)	.672 (.023)	.839 (.026)	.264 (.022)	.690 (.023)	.907 (.020)	.627 (.027)
ROUSE	.763	.909	.454	.727	.840	.190	.745	.948	.603

Note. Within-subject standard error (based on a single-factor, repeated measures analysis of variance): .019 (Loftus & Masson, 1994). ROUSE = responding optimally with unknown sources of evidence; N = neither; T = target; F = foil.

Appendix C

Mean Accuracy (With Standard Errors in parentheses) and Corresponding ROUSE Predictions for the Perceptual Identification Task in Experiment 3

	Condition						
Accuracy	N primed	TA	FA	TI	FI	TI & FA	FI & TA
Observed	.718 (.025)	.684 (.030)	.688 (.031)	.774 (.024)	.639 (.032)	.773 (.022)	.550 (.034)
ROUSE	.723	.698	.703	.765	.642	.756	.528

Note. Within-subject standard error (based on a single-factor, repeated measures analysis of variance): .020 (Loftus & Masson, 1994). ROUSE = responding optimally with unknown sources of evidence; N = neither; TA = target attended; FA = foil attended; TI = target ignored; FI = foil ignored.

Received December 13, 2002
Revision received January 6, 2004
Accepted June 17, 2004 ■