Journal of Experimental Psychology: General
1998, Vol. 127, No. 3, 286–303

0096-3445/98/$3.00

Replicable Unconscious Semantic Priming

Sean C. Draine and Anthony G. Greenwald
University of Washington

In 4 experiments, subjects classified visually presented target words as pleasant–unpleasant words or male–female first names. Prime words were similar (congruent) or dissimilar (incongruent) in meaning to targets. Brief duration of prime words (17, 33, or 50 ms), along with pre- and postmasking, prevented most subjects from perceiving their physical and semantic properties. By constraining response latencies to fall within a response window—a narrow time band that occurred earlier than subjects would ordinarily respond—these experiments consistently produced subliminal priming effects, indicated by greater error rates for incongruent than congruent priming trials. This conclusion was confirmed by analyzing magnitude of priming as a regression function of prime perceptibility using the method of A. G. Greenwald, M. R. Klinger, and E. S. Schuh (1995). The data of each experiment passed their significant-intercept criterion for demonstrating unconscious cognition.

In the past decade, cognitive psychologists have been increasingly willing to discuss theories of unconscious cognition. This receptiveness is associated with a series of methodological advances that have led to an increasing variety of demonstrations of unconscious cognition (e.g., Greenwald, Klinger, & Schuh, 1995; Jacoby, Lindsay, & Toth, 1992; Reingold & Merikle, 1988). Despite these advances, empirical evidence for unconscious perception remains controversial (see review in Greenwald, 1992). Patterns of data that indicate unconscious analysis of "subliminal"[1] stimuli have been only sporadically and weakly producible. Skeptics have been inclined to interpret these empirical inconsistencies as indicating the lack of any valid phenomena of unconscious perception (e.g., Holender, 1986).

This article works toward removing controversy by introducing methods that allow subliminal semantic-priming effects to be easily replicable. The main ingredient of these methods is a response-window procedure, which is used in association with a semantic-priming task. The response-window procedure uses long-established principles of speed–accuracy trade-off (e.g., Reed, 1973; Wickelgren, 1977) to increase the sensitivity of measures of semantic priming.

Background: Using Direct and Indirect Measures to Reveal Unconscious Cognition

Studies of subliminal semantic activation (SSA)—defined as evidence for analysis of a word's meaning under conditions that prevent awareness of the word's physical presence—usually compare instructed responses to masked or otherwise degraded stimuli (direct effects) with uninstructed responses that reveal some evidence of semantic analysis (indirect effects) of those stimuli. Direct effects are usually assessed with measures of the accuracy of judgments of physical characteristics of task stimuli such as presence or absence (Marcel, 1983; Fowler, Wolford, Slade, & Tassinary, 1981) or their spatial position on a display (Greenwald et al., 1995). Indirect effects are often assessed by measuring the influences of masked or unattended task-irrelevant (distractor) stimuli on the latency or accuracy of the instructed response to clearly perceivable task-relevant stimuli. Examples of indirect effects include semantic-priming effects (Meyer & Schvaneveldt, 1971) and the well-studied Stroop (1935) effect.

A widely accepted operational definition of SSA is a pattern of data showing statistically significant indirect effects for stimuli that produce no evidence of direct effects (Dixon, 1971; Greenwald et al., 1995; Holender, 1986). Although there are well-developed arguments for using other empirical criteria (e.g., Cheesman & Merikle, 1984; Greenwald et al., 1995; Jacoby et al., 1992; Reingold & Merikle, 1988), this indirect-without-direct-effect data pattern has remained the most generally accepted operational definition of SSA. Remarkably, despite the wide acceptance of the suitability of the indirect-without-direct-effect pattern and despite multiple published demonstrations of indirect effects without direct effects,[2] SSA has not achieved the

Sean C. Draine and Anthony G. Greenwald, Department of Psychology, University of Washington.

This research was supported by National Institute of Mental Health Grant MH41328 and National Science Foundation Grant SBR-9422242.

Correspondence concerning this article should be addressed to Sean C. Draine or Anthony G. Greenwald, Department of Psychology, University of Washington, Box 351525, Seattle, Washington, 98195–1525. Electronic mail may be sent to seandr@microsoft.com or agg@u.washington.edu.

[1] Although the term *subliminal* implies a long-abandoned theory of the perceptual threshold, or limen, this article nevertheless uses the term, both because it is widely recognized and used by the general (nontechnical) population as a designation for marginally perceptible stimuli and because it continues to be used routinely by psychologists, even those who no longer accept the concept of well-defined thresholds.

[2] In addition to research reviewed in the following section, see Avant and Thieman (1985), Brown and Hagoort (1993), Dagenbach, Carr, and Wilhelmsen (1989), Doyle and Leach (1988),

status of a generally accepted, empirically established phenomenon. The reason for this state of affairs is not difficult to identify. SSA has remained a recalcitrant phenomenon. Many attempts to produce SSA fail, and most published demonstrations report weak effects.

Existing Efforts to Establish the Indirect-Without-Direct-Effect Data Pattern

SSA entered contemporary cognitive psychology when Marcel (1980, 1983) reported findings that showed the indirect-without-direct-effect data pattern. Soon afterwards, a few other investigators (Balota, 1983; Fowler et al., 1981) similarly reported having obtained the indirect-without-direct-effect pattern. However, these findings were immediately controversial for reasons that were well-described by Holender (1986), who pointed out limitations in the sensitivity of the direct measures used in those studies. Furthermore, Cheesman and Merikle (1984) were unable to reproduce the indirect-without-direct-effect data pattern when using direct measures that were more sensitive than those of the earlier studies by Marcel, Fowler et al., and Balota. Consequently, Holender's skeptical appraisal of SSA was well-justified.

Cheesman and Merikle (1984, 1986) offered an alternative interpretation of SSA that required indirect effects to be obtained, not in the absence of direct effects (i.e., not at an objective threshold on the direct measure), but rather when perception of the subliminal stimulus was below a subjective threshold. Cheesman and Merikle described the subjective threshold for a stimulus as the highest level of stimulus presentation at which a subject reports subjective unawareness of its presence. This new criterion was consistent with a reasonable interpretation of unconscious as meaning a lack of subjective awareness.

Not surprisingly, empirical support for SSA was more easily obtainable with the relaxed empirical criterion represented by the subjective threshold notion. Perhaps also not surprisingly, Cheesman and Merikle's (1984) analysis and findings did not undo the prevailing skepticism regarding SSA. An important limitation of the subjective threshold conception was the difficulty of giving it a precise operational definition. From the long tradition of research on signal-detection theory (Green & Swets, 1966), it is well known that the stimulus-presentation conditions at which any perceiver places the boundary between judged presence and absence of a stimulus can be influenced by instructional or motivational variations. It is difficult to accept a subject's assertion of subjective absence of a stimulus at face value when it is known that the subject, with somewhat different instructions, might have indicated presence (for discussion of problems with subjective measures of awareness, see Eriksen, 1960, and Merikle, 1984).

In Cheesman and Merikle's (1984) interpretation, conscious cognition occurred only for stimuli above subjective threshold, implying thereby that direct effects of stimuli below subjective threshold must reflect unconscious cognition. Reingold and Merikle (1988) captured this implication in the form of an *inclusiveness assumption*: that performance on direct measures (and also on indirect measures) can reflect both conscious and unconscious influences. Reingold and Merikle showed that this inclusiveness assumption, along with a relative sensitivity assumption—that the direct measure is at least as sensitive as the indirect measure to conscious perception of relevant stimulus information—allowed an indirect-greater-than-direct-effect pattern (of which the indirect-without-direct-effect pattern is a special case) to provide evidence for unconscious cognition.

The Reingold and Merikle (1988) analysis, as described previously, appeared to increase greatly the possibilities for obtaining findings that could support SSA. However, the indirect-greater-than-direct pattern has proven difficult to obtain. It has appeared most notably in research using measures of direct and indirect effects obtained at a substantial time delay after original stimulus presentations (Bonanno & Stillings, 1986; Bornstein & D'Agostino, 1992; Kunst-Wilson & Zajonc, 1980; Mandler, Nakamura, & Van Zandt, 1987; Merikle & Reingold, 1991; Seamon, Marsh, & Brody, 1984). It has not been obtained in semantic-priming experiments.

Necessarily, the indirect-greater-than-direct pattern indicates unconscious cognition occurring in association with conscious cognition—not in the absence of conscious cognition. The question of whether unconscious cognition can occur in the absence of conscious cognition was set to the side by Reingold and Merikle's (1988) analysis. Nevertheless, it is the latter form of unconscious cognition—evidence of semantic activation in the absence of conscious cognition—that remains the focus of greatest interest and controversy, and it is also that form of unconscious cognition that is central to the conceptualization of SSA. Convincing evidence for SSA still requires the indirect-without-direct-effect data pattern.

Regression Strategy for Identifying the Indirect-Without-Direct-Effect Pattern

One of the main difficulties of demonstrating the indirect-without-direct pattern is that it is necessary to claim that one has evidence in favor of a null hypothesis: specifically, the null hypothesis that there is no direct effect of the subliminally presented stimulus. Such an assertion of truth of the null hypothesis is statistically insecure because standard statistical methods describe the unacceptability of a null hypothesis, not its acceptability.

In order to overcome the statistical problem of needing to accept a null hypothesis, Greenwald et al. (1995) introduced a regression-analysis method in which the indirect measure, treated as criterion, is examined as a linear-regression function of a direct measure, treated as predictor. The regression analysis yields a slope estimate (b) that measures the extent of association between direct and indirect measures. More important, if direct and indirect measures are assessed on scales that have rational zero points, the regression yields an intercept estimate (a) that tests the indirect-without-direct pattern. With this regression method,

Greenwald, Klinger, and Liu (1989), Groeger (1988), Hirshman and Durante (1992), Kostandov (1985), and Shevrin (1988).

evidence for SSA takes the form of statistical significance for the test of a positive intercept's difference (D) from zero. A significantly positive intercept means that above-zero performance on the indirect measure is associated with null (zero) performance on the direct measure. Consequently, the statistical test that confirms the indirect-without-direct pattern is a null hypothesis rejection—finding that the intercept estimate is significantly greater than zero—rather than a null hypothesis acceptance.

Using the regression method, Greenwald et al. (1995) and Greenwald and Draine (1997) found evidence for SSA in the form of statistically significant intercept effects. Although the statistical reliability of effects in each of these studies was not in question, the intercept effects were nevertheless numerically small and were not obtained consistently in all experimental tests. In terms of Cohen's (1988) effect size measure **d**, the intercepts fell within the range of $\mathbf{d} = 0.1$ to $\mathbf{d} = .2$, a level that is conventionally regarded as small. In summary of these initial tests of SSA with the regression method: The method was successfully demonstrated, but procedures that can be counted on to replicably produce substantial SSA effects were not demonstrated.

Development of the Response-Window Method

Dilution of Priming Effects Across Latency and Accuracy Measures

As noted by Wickelgren (1977) and others, the possibility of subjects trading off speed for accuracy (becoming more accurate when responding more slowly) complicates the interpretation of effects of independent variables on response latencies. Wickelgren's analysis was primarily concerned with the potential for trade-offs to result in Type I errors by producing latency differences between experimental conditions that were equal in cognitive difficulty. Wickelgren's arguments also imply that speed–accuracy trade-offs can lead to Type II errors by obscuring potential latency differences across conditions that differ in the cognitive demands imposed upon subjects.

Consider a typical priming procedure in which the subject is instructed to respond rapidly to target stimuli while making as few errors as possible. In this procedure, priming effects are expected to emerge as shorter response latencies to targets preceded by semantically related (or congruent) than by semantically unrelated (or incongruent) primes. By focusing on latencies as the primary dependent variable, researchers tacitly assumed that subjects perform this task by withholding their response to a target until the total accumulated target information exceeds a self-imposed criterion associated with high accuracy (Grice, Nullmeyer, & Spiker, 1977). According to this view, semantic-priming effects should be measurable as differences in response latencies across prime conditions. On the other hand, subjects may perform priming tasks by setting an internal time criterion for responding that is relaxed (i.e., long) enough to yield a low overall error rate (Ollman, 1977). For example, subjects may fall into a steady rhythm of responding due to the repetitive nature of many priming tasks. In this case priming effects may be more readily measurable by differences in accuracy rather than latency.

If subjects vary in whether they base their responses on a latency or an accuracy criterion, or if the same subject uses both of these strategies during an experiment, then priming effects will be distributed across measures of accuracy and latency. Further, if priming becomes more evident in one dimension only by becoming less evident in the other, then magnitude of priming effects measured separately in both dimensions is smaller than if the priming effect were concentrated on either dimension alone. The dilution of priming effects across speed and accuracy measures may be especially problematic when the accuracy measure is near ceiling (i.e., very low error rates), as is encouraged by typical instructions for speeded responding tasks. In this situation, the part of the priming effect that might be evident as differences in accuracy becomes difficult to observe.

The Response-Window Procedure

In order to manage the difficulties caused by possible dilution of priming effects across latency and accuracy measures, one might attempt to combine the two measures into a single index (e.g., bits of information transmitted per second; see Attneave, 1959). However, finding an appropriate mathematical formula for combining accuracy and latency measures depends on knowing the speed–accuracy trade-off function. Consequently, this strategy cannot be used effectively when the speed–accuracy trade-off function is unknown. A second strategy, used by Dosher and Rosedale (1991), is to vary response latencies through the use of response cues or deadlines in order to obtain enough empirical points on the speed–accuracy plane to estimate the speed–accuracy trade-off curve. A disadvantage of this strategy is that it requires observation of a very large number of trials.

A third strategy, developed in these experiments, is to oblige subjects to hold one dimension of performance constant so that any effects of the independent variable are confined to the other dimension. Subjects could in principle be encouraged to hold either their error rates or latencies constant. The strategy of controlling error rates is indeed typically used in reaction time tasks, by requesting subjects to maintain errors at a low level. However, the subject's cooperation with this request does not assure that error rates are indeed constant across conditions. Further (as observed by Wickelgren, 1977, and others) when error rates are low, floor effects make it difficult to observe possibly important between-condition differences in accuracy. By contrast, it is comparatively easy to control latencies by providing an external timing signal to facilitate the subject's maintenance of a regular rhythm of responding.

These experiments controlled response latencies by instructing subjects to respond within an interval of time that can be described as a response window. The response window is defined by two parameters: window center (the center of the latency range to which the subject is trying to

conform) and the window width (the width of the targeted latency range). In this research, these window parameters were set at values that produced error rates substantially higher than are characteristic of previous SSA research.

Experiment 1

Experiment 1 sought to determine whether the response-window procedure would increase the power of the regression method for detecting unconscious cognition. The experiment used two tasks, the first providing an indirect measure of semantic priming and the second a direct measure of perceptual discriminability of the primes. Both semantic priming and prime perceptibility were measured with visually masked prime stimuli presented at three durations: 17, 33, and 50 ms.

Method

Subjects

Subjects were 43 undergraduate students at the University of Washington (Seattle) who volunteered to participate in the experiment in exchange for extra credit in an introductory psychology course. All were self-described as fluent in English and as having normal or corrected-to-normal vision.

Apparatus

Up to 3 subjects participated concurrently, each in a separate cubicle with a 33-cm (diagonal) color monitor and keyboard controlled by an IBM/AT-type (80486) computer. A fan motor in each cubicle produced background white noise to mask extraneous sounds. Subjects viewed a color (enhanced graphics adapter [EGA]) display through a viewing apparatus that presented the images from the left and right halves of the display screen to the left and right eyes, respectively. (The same type of apparatus was used by Cheesman & Merikle, 1986; Greenwald & Klinger, 1990; and Greenwald et al., 1995.) Subjects viewed the computer's display from a distance of 65 cm through rotary prisms that were adjusted to superimpose the left-eye and right-eye images. All stimuli were presented simultaneously to both halves of the screen and were easily viewed with binocular fusion. (This dichoptic viewing arrangement was not essential to this research, but it was used because it had been found to increase masking effectiveness modestly relative to ordinary binocular viewing of the computer display.) Subjects responded on all experimental tasks by pressing either the standard keyboard's *A* key with left forefinger or the *5* key (on the keyboard's right-side numeric keypad) with right forefinger.

Indirect (Priming) Measure: Semantic Classification

In all of the following experiments, indirect measures of semantic priming were obtained using two versions of a semantic-classification task, one involving evaluative-classification judgments and the other gender classification. In the evaluative-classification task, subjects judged whether evaluatively polarized target words were pleasant or unpleasant in meaning. Evaluative stimuli were drawn from the list compiled by Bellezza, Greenwald, and Banaji (1986) by selecting 25 words from the low end of the distribution of normative pleasantness ratings (ranging from 1.28 to 1.61 on a 5-point scale) to serve as unpleasant word stimuli and 25 words from the high end (ranging from 4.43 to 4.83) to serve as pleasant word stimuli. Only words from four to eight letters long and pronounceable in one or two syllables were selected. In the gender-classification task, subjects judged whether targets words were female or male first names. Gender stimuli consisted of 25 male and 25 female first names taken from lists of the names most frequently given to newborns in the United States in 1982 and 1983 (Dunkling & Gosling, 1984). All names were from three to eight letters in length and were pronounceable in one or two syllables.

On both versions of the semantic-classification task, a string of 13 uppercase consonant letters (e.g., KQHYTPDQFPBYL) marked the beginning of the trial and also served as a forward mask. This forward mask remained on the screen for 150 ms and was immediately replaced by a prime word, displayed in uppercase letters, that was either evaluatively polarized in the evaluative-classification task or a male or female first name in the gender-classification task. Prime duration was a within-block variation, randomly varying from trial to trial among the values of 17, 33, and 50 ms using an on-line randomization algorithm that (a) presented each prime stimulus equally often at each prime duration and (b) conducted an equal number of trials for each prime-duration condition. The prime words were replaced by a different string of 13 uppercase consonant letters that served as a backward mask, which was displayed for 17 ms. On all trials, a clearly visible target word, displayed in lowercase letters, was presented 67 ms following the onset of the prime word and remained on the screen for 183 ms. The 67-ms interval between prime and target onsets defined the stimulus onset asynchrony (SOA) of the priming task. Targets were evaluatively polarized words in the evaluative-classification task or a male or female first name in the gender-classification task. Finally, the target word was replaced after 183 ms by an exclamation point that defined the response window. The window center was 250 ms following target onset, and the window width was 133 ms. (The exclamation point was thus on screen from 183 to 317 ms after the target word onset.) All of the stimuli were presented in fixed-width black letters on a strip of light gray background (1 character space tall and 21 character spaces wide) located in the middle of the otherwise dark gray screen.

The complete list of stimuli is given in the appendix. Stimuli were randomly selected so that (a) each appeared exactly once as the target and once as the prime within each block of 50 trials and (b) any one stimulus did not appear as both prime and target on the same trial. On each trial of the semantic-classification task, the relation between prime and target was either congruent or incongruent. Congruent trials were those in which the prime and target were both pleasant or both unpleasant for evaluative classifications and both male or both female for gender classifications. Incongruent trials were those in which one of the prime–target pair was pleasant and the other unpleasant for evaluative classifications, and one of the pair was male and the other female for gender classifications. During practice blocks, primes and targets were always congruent. During test blocks, primes and targets were congruent on a randomly selected 50% of the trials and incongruent on the remaining 50% of trials.

Subjects were instructed to ignore the mask and prime stimuli and to classify the target words as either pleasant or unpleasant in the evaluative-classification task or as either male or female in the gender-classification task. Subjects classified targets as unpleasant (male) or pleasant (female) by pressing the left or right response key, respectively, in the evaluative- (gender-) classification task. For blocks of trials using the response window, subjects were instructed to respond while the black exclamation point was on the screen. Feedback for success in responding during the window interval was provided by the exclamation point's behavior. If the response was early, the exclamation point never appeared on the

screen. If the response occurred during the window interval, the exclamation point changed from black to red and remained on the screen for the remaining 300 ms of the trial. If the exclamation point appeared and disappeared without changing color, the subject knew that the response was late. The next trial started (with the 13-character premasking consonant string) 600 ms after the subject pressed a response key.

Direct (Prime Perceptibility) Measure: Word Versus XG-String Discrimination

Experiment 1 included a second task that provided a direct measure of the perceptibility of the prime words used in the semantic-classification task. In the direct-measure task, subjects were asked to judge whether the stimulus appearing between the forward and backward mask was a word or a string of alternating Xs and Gs. Word and mask stimuli were identical to those used in the semantic-classification task. XG-string stimuli were constructed from the evaluative and gender stimuli by replacing each character in those words and names alternately with an X or a G (e.g., the replacement for EVIL could be GXGX or XGXG). This procedure produced two lists of 50 XG strings, one for evaluative stimuli and one for gender stimuli, with each string identical in length to its corresponding word stimulus. None of the mask, evaluative-word, or first-name stimuli contained either an X or a G.

In the same fashion as for the indirect-measure (priming) task, the stimulus sequence for each trial of the prime perceptibility task began with a string of 13 uppercase consonants serving as a forward mask that remained on the screen for 150 ms. The forward mask was immediately replaced by an uppercase letter string that served as the critical stimulus for this task. As for the indirect-measure task, critical stimulus duration was a within-block variation, randomly varying from trial to trial among the values of 17, 33, and 50 ms such that there were equal numbers of trials for each prime duration. Critical stimuli were randomly selected from the words or names used on indirect measures and the set of XG strings. The selection algorithm ensured that words (or names) and XG strings appeared equally often and that each stimulus was presented equally often at each duration. The critical stimulus was immediately replaced by a backward mask that lasted 17 ms. The backward mask was followed, after a blank period (if needed) to fill out the 67-ms (SOA), by a distractor word displayed in lowercase. Distractor words were randomly selected on each trial from the set of word (or name) stimuli used on indirect measures, with the same stimulus never appearing as both critical stimulus and distractor on the same trial. Although the distractor words were irrelevant to the task, they were included to ensure that the stimulus-presentation sequence of the direct measure's prime-perception task matched that of the indirect measure's semantic-classification task. (Any influence of the target stimuli on the visibility of prime stimuli should therefore affect direct and indirect measures equally.)

Subjects were instructed to ignore the clearly visible distractor words and to classify the preceding masked word as being either a string of alternating Xs and Gs or as a word or name. Subjects classified the stimulus as a word or name by pressing the right response key and as an XG string by pressing the left response key. Subjects were instructed to respond as quickly and as accurately as possible. No response window was used in this task because of preliminary findings showing that inclusion of the window impaired direct-measure performance.[3] Six hundred milliseconds after subjects responded, the next trial began with its 13-character premasking consonant string.

Comparability of Direct and Indirect Measures

Because indirect (priming) effects are driven by the semantic category of the primes, the relative sensitivity assumption (cf. Greenwald et al., 1995; Reingold & Merikle, 1988) requires that the direct measure chosen as the predictor in the regression analysis be at least as sensitive as the indirect measure to conscious perception of the prime's semantic category. The XG string versus word discrimination (XGvWd hereafter) task, which requires subjects to make perceptual discriminations of physical properties of the primes, was chosen because it seemed likely to meet this criterion. First, it was assumed that subjects who could consciously perceive any semantic feature of the prime should thereby be able to perform the XGvWd discrimination. Second, subjects had an additional means of performing the XGvWd task accurately if they could consciously perceive enough physical stimulus information to make out even so much as a single character. For these reasons, the XGvWd task was considered, a priori, to have the desired characteristics of being at least as sensitive (and likely more sensitive) to consciously perceived stimulus information as the indirect measure.

Procedure

Semantic-classification (indirect-measure) task: Practice. The first task for all subjects was semantic classification. Half of the subjects performed the evaluative-classification version and the remainder performed the gender-classification version. Each subject performed initially a minimum of 200 practice trials. The first block of 20 practice trials contained only target words (no masks or primes). Next, subjects performed two blocks of 50 trials in which masked primes were presented before the targets, with prime duration fixed at 50 ms. On all of these trials, subjects received immediate feedback in the form of the displayed word ERROR when they incorrectly classified the target word. At the end of each of these blocks, subjects were informed of their percentage of correct responses for that block and were given the option to repeat the block or proceed to the next phase of the experiment. The next block of 20 practice trials introduced the response-window procedure. On the first 5 trials of this block and also of all subsequent blocks in which the response-window procedure was used, subjects were instructed to observe the sequence of stimuli without giving a response so that they could learn the placement of the window interval. Subjects responded on the remaining trials of the block and were given feedback (described previously) about success in responding during the window interval. Subjects had the option of repeating these 20 practice trials multiple times if desired before performing a final block of 50 practice trials with the response window.

On all response-window trials, the ERROR message did not appear after incorrect responses. At the end of all blocks of practice and test trials involving the response window, however, subjects

[3]Specifically, a response-window procedure was found to greatly reduce accuracy on a direct measure for which subjects classified masked stimuli that were presented at durations of 17, 33, 50, and 100 ms as words or strings of digits. If, as many have argued (e.g., Posner & Snyder, 1975; Fodor, 1983), conscious cognition operates more slowly than unconscious cognition, the response-window procedure (which forces rapid responding) should reduce sensitivity more to conscious than to unconscious cognitive influences. Use of the response-window procedure with direct measures would therefore undermine the goal of using direct measures that are maximally sensitive to conscious stimulus effects.

learned their percentage success in responding during the window and were encouraged to keep this percentage at 70% or higher. Additionally, they learned their percentage of correct classifications and were advised that although relatively high error rates could normally occur, they should nevertheless try to respond as accurately as possible.

Semantic-classification task: Data collection. After completion of practice, subjects began the data collection phase of the semantic-classification task, which consisted of six blocks of 50 trials (300 trials total). As described previously, a randomly selected 50% of these trials consisted of congruent prime–target pairs, with the remaining trials consisting of incongruent pairs. A second independent variable that was varied within blocks was the duration of the prime words. Prime duration varied among values of 17, 33, or 50 ms, determined by an on-line randomization routine that provided equal numbers of trials at each of the three prime durations.

Prime-perception (direct-measure) task: Practice. After the six data collection blocks of semantic classification, subjects proceeded to the XGvWd task. The word stimuli for this task were evaluative words if the subject had previously performed the evaluative version of semantic classification and first names if the subject's previous task had been gender classification. An initial block of 30 practice trials presented clearly visible, unmasked critical stimuli (50-ms duration) in red uppercase letters followed by distractors in black lowercase letters. The different colors and cases were used in conjunction with instructions to help subjects locate, in the stimulus sequence of each trial, which stimulus they were to judge. Subjects then received a second block of 30 practice trials in which the critical stimuli, still displayed in red and still 50 ms in duration, were sandwiched between forward and backward masks consisting of the usual black uppercase 13-character consonant strings. (The red stimuli were usually easily discriminable from the black masks.) On a final block of 30 practice trials and on all subsequent test trials, all stimuli were displayed as black characters. On all practice trials, subjects received immediate feedback in the form of the displayed word ERROR for incorrect classifications. After each practice block, subjects learned their percentage of correct responses for the block and were given the option to repeat the block before proceeding.

Prime-perception task: Data collection. Data collection for the direct-measure task consisted of six blocks of 50 trials. As for the semantic-classification task, the exposure duration of the critical stimuli varied randomly, within blocks, among the values of 17, 33, and 50 ms. All stimuli were selected randomly with the constraints that (a) each of the 50 words (or names) and 50 XG strings appeared once as the to-be-categorized stimulus in each consecutive set of 100 trials (two blocks) and (b) no word could be both the masked critical stimulus and the distractor word on the same trial. On test trials, the ERROR message after incorrect responses was discontinued, although information about overall percentage of correct responses for the block was maintained.

Results

Computation of Direct and Indirect Measures

Data for the two tasks were analyzed using the method of regressing indirect onto direct measures developed by Greenwald et al. (1995). In order to use a common unit for both measures and to provide measures with needed rational zero points, both the indirect measure of priming obtained from the semantic-classification task and the direct measure obtained from the XGvWd task were computed as signal-detection sensitivity (d') measures. For the direct-effect measure, hits were defined as correct classifications of the masked critical stimulus as a word (or name), and false alarms were incorrect classifications of word (or name) in response to masked XG strings. For the indirect measure, hits were judgments of pleasant or female (for evaluative and gender classification, respectively) on trials with pleasant or female primes, and false alarms were judgments of pleasant or female on trials with unpleasant or male primes.

Regression Results

The regression analyses for each prime-duration condition are plotted in Figure 1. The regression analyses are of the combined data from evaluative and gender classifications.[4] With primes presented for 17-ms durations there was evidence for no more than very weak priming, mean $d' = 0.03$, $t(42) = 0.78$, $p = .44$. Estimated intercept, $a = 0.02$, $t(41) = 0.71$, $p = .48$, and slope, $b = 0.01$, $t(41) = 0.20$, $p = .84$, parameters did not differ significantly from zero and corresponded to a regression line that was virtually superimposed on the abscissa. A substantially stronger priming effect was obtained for the 33-ms prime duration, mean $d' = 0.24$, $t(42) = 5.00$, $p = .00001$. For the 33-ms prime duration, the regression slope was flat, $b = -0.02$, $t(41) = -0.17$, $p = .86$, and the intercept was positive, ($a = 0.25$, $t(41) = 3.76$, $p = .0005$. Strong priming effects were also obtained with 50-ms prime durations, mean $d' = 0.21$, $t(42) = 3.99$, $p = .0003$. The regression function for the 50-ms prime yielded a flat slope, $b = 0.05$, $t(41) = 0.48$, $p = .64$, and a positive intercept, $a = 0.18$, $t(41) = 2.57$, $p = .014$. Expressed in terms of Cohen's **d**, the intercept effects for 33-ms and 50-ms prime durations were, respectively, $\mathbf{d} = 0.79$ and $\mathbf{d} = 0.54$, falling approximately at the values conventionally regarded as large and medium, respectively.

To the extent that subjects varied in the amount of time required to process and classify the targets, it follows that some subjects had more difficulty than others responding within the response window. Post hoc analyses of Experiment 1 suggested that the 250-ms window center required responding so rapidly as to reduce overall accuracy levels to the floor (chance) for some subjects, possibly preventing the appearance of priming effects that might otherwise have been evident. Correlations between priming magnitude and overall accuracy levels were consistent with this suspicion. Subjects with higher overall error-rate levels showed less priming in both the 33-ms, $r(42) = -.56$, $N = 43$, $p = .0001$, and 50-ms, $r(42) = -.47$, $N = 43$, $p = .002$, prime-duration conditions.

[4] Separate regression analyses of gender and evaluative classifications were conducted on the aggregated data sets of Experiments 2–4 that are reported later. Observed regression slopes and intercepts generally were quite similar for evaluative and gender classification, justifying the procedure of combining the data from the two tasks for analyses of the individual experiments.

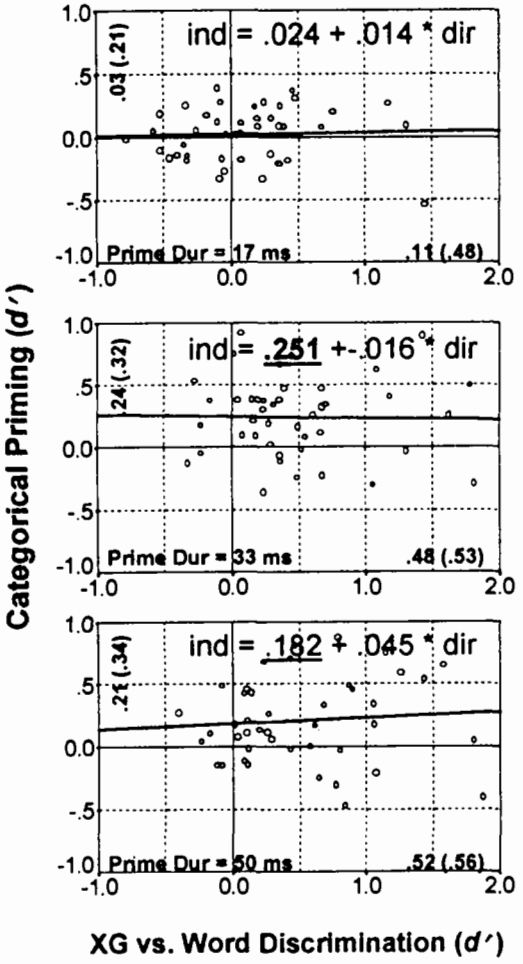

Figure 1. Indirect effects regressed on direct effects, separately for the 17- (top), 33- (middle), and 50-ms (bottom) prime-duration (Prime Dur) conditions of Experiment 1 ($N = 43$). Data points are individual subjects. Evaluative- and gender-classification conditions are represented indistinguishably in the plots. The direct measure (dir) of stimulus perceptibility is represented on the abscissa, and the indirect measure (ind) of semantic priming is on the ordinate. Means (and standard deviations in parentheses) for direct and indirect measures are shown in the lower right and upper left corners of the plots, respectively. Equations for best fitting linear-regression functions appear at the top of each scatterplot. In these equations, statistically significant intercept and slope parameters are underlined ($p < .05$); bold and underlined ($p < .005$); or bold, underlined, and italic ($p < .0005$). The asterisk (*) indicates multiplication. XG = XG-string stimuli that were constructed from the evaluative and gender stimuli by replacing each character in those words and names alternately with an X or a G (e.g., the replacement for EVIL could be GXGX or XGXG).

Experiment 2

The purpose of Experiment 2 was to replicate Experiment 1 using a response-window procedure that would allow, for each subject, performance at the indirect measure's semantic-classification task that was at least moderately above chance accuracy. Therefore, an adaptive response-window procedure was used in which the window center was adjusted contingent on the subject's accuracy. The window center was thus tailored to each subjects' performance and could also be adjusted with changes in the subject's performance accuracy during the experiment.

Method

Subjects

Subjects were 38 undergraduate students at the University of Washington (Seattle) who volunteered to participate in exchange for extra credit in an introductory psychology course. All were fluent in English and had normal or corrected-to-normal vision. Data from one subject who did not complete the experiment because of eye fatigue were excluded from all of the analyses reported later.

Procedure

The procedure used in Experiment 2 was in most respects identical to that of Experiment 1. One exception was that the number of test trials for the semantic-classification task was increased to a total of 450 (nine blocks of 50 trials) in order to boost power. Second, an adaptive rather than fixed response window was used. For all subjects, the response window was initially centered at 400 ms after onset of the target word. At the end of each block, the window center could be made shorter by 33 ms, longer by 33 ms, or could remain unchanged, depending upon the subject's performance in that block. The window center was made shorter if the subject's error percentage was less than or equal to 20% and the subject's mean response latency for that block was no more than 100 ms greater than the current window center. The window center was made longer if the subject's error percentage was greater than or equal to 45% and the subject's mean response latency was more than 100 ms longer than the current window center. If neither of these sets of conditions was met, the window center was not changed.

Results and Discussion

Figure 2 shows the results of regression analyses for Experiment 2. As in Experiment 1, no statistically significant evidence for priming was obtained for 17-ms primes, mean $d' = 0.02$, $t(36) = 0.57$, $p = .57$. The regression analysis for that condition yielded a slightly positive intercept, $a = 0.02$, $t(36) = 0.46$, $p = .65$, and a nearly flat slope, $b = 0.03$, $t(36) = 0.32$, $p = .75$. For 33-ms primes, the overall priming effect was substantial, mean $d' = 0.35$, $t(36) = 7.19$, $p = 10^{-8}$, and the regression analysis yielded a strong intercept, $a = 0.33$, $t(36) = 5.84$, $p = 10^{-6}$, and a nearly flat slope, $b = 0.04$, $t(36) = 0.47$, $p = .64$. For 50-ms primes, overall priming was also substantial, mean $d' = 0.50$, $t(36) = 9.59$, $p = 10^{-11}$, the regression intercept was sizable, $a = 0.44$, $t(36) = 7.14$, $p = 10^{-8}$, and the slope was positive but nonsignificant, $b = 0.11$, $t(36) = 1.56$, $p = .13$.

The finding of larger priming in Experiment 2 than in Experiment 1 suggested that priming effects were captured more efficiently by Experiment 2's adaptive window procedure than by Experiment 1's fixed 250-ms window center. Overall error rates with the adaptive window procedure were weakly and nonsignificantly correlated with priming magni-

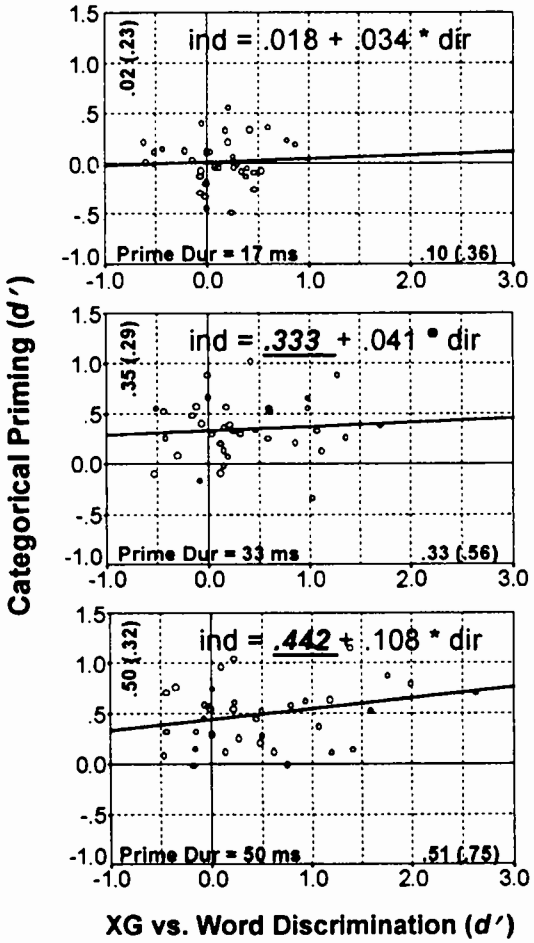

Figure 2. Scatterplots and best fitting linear-regression functions separately for the 17- (top), 33- (middle), and 50-ms (bottom) prime-duration (Prime Dur) conditions in Experiment 2 (N = 37). Data points are individual subjects. Evaluative- and gender-classification conditions are represented indistinguishably in the plots. The direct measure (dir) of stimulus perceptibility is represented on the abscissa, and the indirect measure (ind) of semantic priming is on the ordinate. Means (and standard deviations in parentheses) for direct and indirect measures are shown in the lower right and upper left corners of the plots, respectively. Equations for best fitting linear-regression functions appear at the top of each scatterplot. In these equations, statistically significant intercept and slope parameters are printed in underlined ($p < .05$); bold and underlined ($p < .005$); or bold, underlined, and italic ($p < .0005$). The asterisk (*) indicates multiplication. XG = XG-string stimuli that were constructed from the evaluative and gender stimuli by replacing each character in those words and names alternately with an X or a G (e.g., the replacement for EVIL could be GXGX or XGXG).

tude in the 33-ms, $r(36) = -.25$, $N = 37$, $p = .13$, and 50-ms, $r(36) = -.07$, $N = 37$, $p = .68$, prime-duration conditions, in contrast with the stronger correlations of priming with error rates observed in Experiment 1. This observation supports the speculation that the fixed 250-ms window center of Experiment 1 stifled priming in some subjects by forcing nearly chance responding on the classification task used for the indirect measure.

Experiment 3

In Experiments 1 and 2, variations of prime duration were confounded with variations of the interstimulus interval (ISI): that is, the interval between offset of the prime and onset of the target. In Experiments 1 and 2, shorter prime durations were associated with longer ISIs: The ISI for 50-ms primes was 17 ms (the length of the postmask), the ISI for 33-ms primes was 33 ms, and the ISI for 17-ms primes was 50 ms. If the influence of unconsciously processed primes decreases as a function of interval after termination of the prime stimulus, this lengthening of the ISI would be responsible for reduced priming.

The purpose of Experiment 3 was to separate effects on priming attributable to prime duration from those due to ISI. Primes of 17-ms duration were presented with ISIs of either 17 or 50 ms, and primes of 33-ms duration were presented with ISIs of 17 or 33 ms.

Method

Subjects

Subjects were 48 undergraduate students at the University of Washington (Seattle) who volunteered to participate in the experiment in exchange for extra credit in an introductory psychology course. All were fluent in English and had normal or corrected-to-normal vision. Data from 4 of these subjects were lost because of equipment malfunction. In addition, 1 subject was unable to complete the experiment in the scheduled 1-hr time slot, and another subject was unable to complete the experiment because of eye fatigue. Data from these 2 subjects were excluded from all analyses reported later.

Procedure

The procedures of Experiment 3 were similar to those of Experiment 2, incorporating the two independent variables of prime duration and ISI. There were four conditions, which can be identified by their prime duration–ISI values (in milliseconds) as follows: 17/17, 17/50, 33/17, and 33/33. (SOAs in each case were necessarily the sum of the two values shown.)

The response-window procedure made it unfeasible to vary the four conditions from trial to trial as had been done with prime duration in the previous experiments. This was because variations in ISI produced variations in the length of time between the warning stimulus (i.e., the forward mask) and the to-be-judged target stimulus. Trial-to-trial variation of ISI would have disrupted subjects' efforts to coordinate responses with the window interval because the interval from trial start to window center would have differed from trial to trial. Therefore, the four conditions were manipulated randomly on a block-by-block basis. Each subject first performed 12 blocks of 50 trials of the semantic-classification task and then 8 blocks of the XGvWd task. For both indirect and direct measures, blocks were distributed equally among the four conditions, resulting in a total of 135 trials per condition for indirect measures (no response was given on the first 5 trials of each block) and 100 trials per condition for direct measures.

Results and Discussion

Regression Results

Experiment 3 yielded evidence of unconscious semantic priming with a prime duration of 17 ms. The regression analysis for the 17/17 condition shown in Figure 3 revealed a positive intercept, $a = 0.10$, $t(40) = 2.88$, $p = .006$, and a flat slope, $b = -0.03$, $t(40) = -0.23$, $p = .82$. For the 17/50 condition, the regression analysis also yielded a positive intercept, $a = 0.09$, $t(40) = 2.99$, $p = .005$, and a weakly positive slope, $b = 0.13$, $t(40) = 1.16$, $p = .26$. Regression analysis of the 33/17 condition shown in Figure 4 revealed a positive intercept, $a = 0.16$, $t(40) = 3.71$, $p = .0006$, and a flat slope, $b = 0.04$, t(40) = 0.38, $p = .71$. Lastly, for the 33/33 condition, the regression analysis also yielded a positive intercept, $a = 0.34$, $t(40) = 8.75$, $p = 10^{-8}$, and, again, a flat slope, $b = -0.01$, $t(40) = -0.19$, $p = .85$.

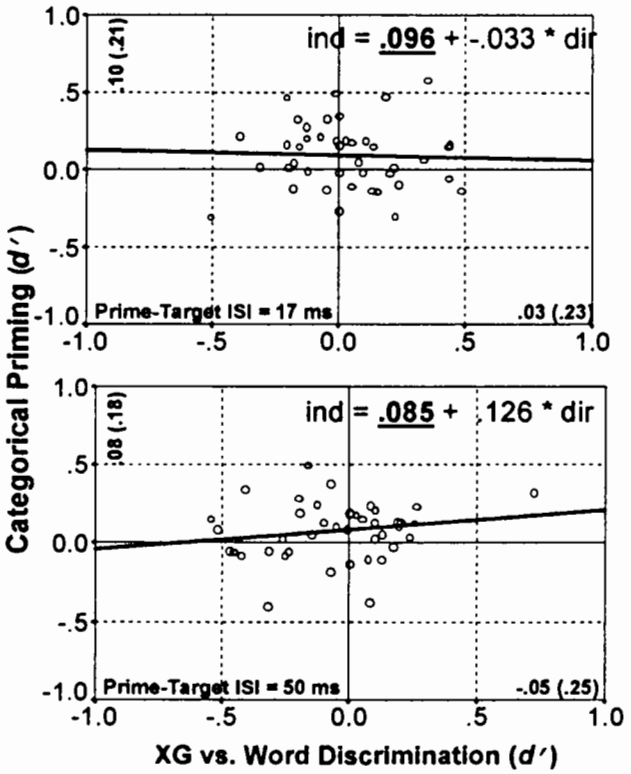

Figure 3. Scatterplots and linear-regression functions for the 17-ms prime duration and 17-ms prime–target interstimuls interval (ISI; top) and the 17-ms prime duration and 50-ms ISI (bottom) conditions of Experiment 3 ($N = 42$). Data points are individual subjects. Evaluative- and gender-classification conditions are represented indistinguishably in the plots. The direct measure (dir) of stimulus perceptibility is represented on the abscissa, and the indirect measure (ind) of semantic priming is on the ordinate. Means (and standard deviations in parentheses) for direct and indirect measures are shown in the lower right and upper left corners of the plots, respectively. Equations for best fitting linear-regression functions appear at the top of each scatterplot. In these equations, statistically significant intercept and slope parameters are printed in underlined ($p < .05$); bold and underlined ($p < .005$); or bold, underlined, and italic ($p < .0005$). The asterisk (*) indicates multiplication. XG = XG-string stimuli that were constructed from the evaluative and gender stimuli by replacing each character in those words and names alternately with an X or a G (e.g., the replacement for EVIL could be GXGX or XGXG).

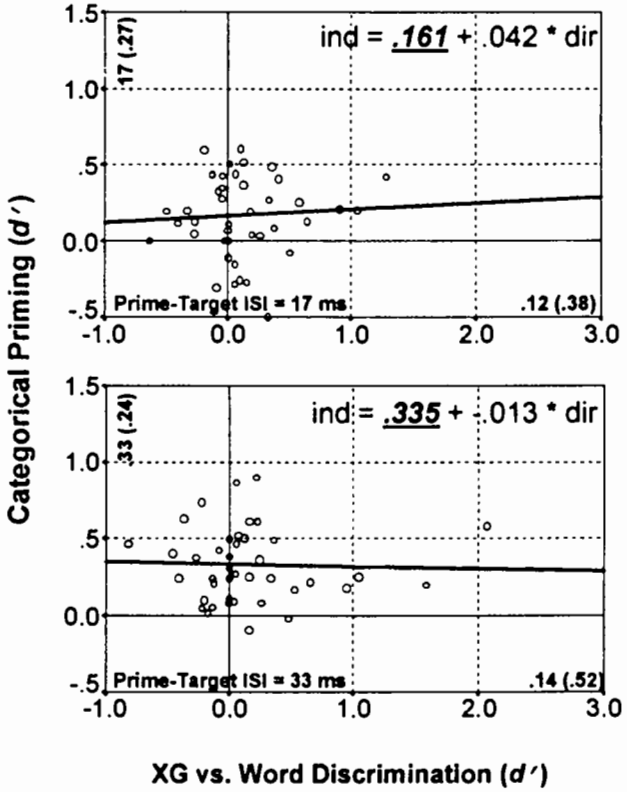

Figure 4. Scatterplots and linear-regression functions for the 33-ms prime duration and 17-ms interstimulus interval (ISI; top) and the 33-ms prime duration and 33-ms ISI (bottom) conditions of Experiment 3 ($N = 42$). Data points are individual subjects. Evaluative- and gender-classification conditions are represented indistinguishably in the plots. The direct measure (dir) of stimulus perceptibility is represented on the abscissa, and the indirect measure (ind) of semantic priming is on the ordinate. Means (and standard deviations in parentheses) for direct and indirect measures are shown in the lower right and upper left corners of the plots, respectively. Equations for best fitting linear-regression functions appear at the top of each scatterplot. In these equations, statistically significant intercept and slope parameters are printed in underlined ($p < .05$); bold and underlined ($p < .005$); or bold, underlined, and italic ($p < .0005$). The asterisk (*) indicates multiplication. XG = XG-string stimuli that were constructed from the evaluative and gender stimuli by replacing each character in those words and names alternately with an X or a G (e.g., the replacement for EVIL could be GXGX or XGXG).

Effect of Prime–Target ISI on Priming Magnitude

The mean priming effects for the four combinations of prime duration and ISI are shown in the upper left corner of the corresponding panels in Figures 3 and 4. Statistically significant priming effects were obtained for 17-ms primes

with both the 17-ms prime–target ISI, $t(41) = 2.91$, $p = .006$, and the 50-ms prime–target ISI, $t(41) = 2.80$, $p = .008$. The difference between these two priming effects was quite small and not statistically significant, $D = .02$, $t(41) = 0.41$, $p = .69$. For the 33-ms prime duration, statistically significant priming was obtained with both the 17-ms prime–target ISI, $t(41) = 4.04$, $p = .0002$, and the 33-ms prime–target ISI, $t(41) = 9.15$, $p = 10^{-11}$. In contrast to expectations, priming by 33-ms duration primes was greater with the longer rather than the shorter ISI condition, $D = -.17$, $t(41) = -3.58$, $p = .0009$.

In sum, the results of Experiment 3 did not support the hypothesis that longer ISIs result in smaller priming effects. The chief implication is for interpretation of results of the preceding two experiments. In particular, the relatively weak priming found with 17-ms prime durations in the previous two experiments is better interpreted as the result of shorter prime duration rather than longer ISI (i.e., the interval between end of prime and start of target).

Experiment 4

Experiments 1–3 have established that the response-window method yields replicable semantic-activation effects of stimuli presented at the margin of visual perceptibility. Experiment 4, the last in this series, was designed to confirm that these subliminal semantic activation (SSA) effects warrant interpretation as effects of unconscious cognition and possibly as effects that are dissociated from conscious cognition.

The intercept effects obtained in Experiments 1–3 describe the mean priming effect associated with primes that were perceptually indistinguishable from strings of Xs and Gs. If the direct and indirect measures used in these studies meet the regression method's assumptions (see Greenwald & Draine, 1997, p. 100; Greenwald et al., 1995, p. 26), then the intercept effects can be further interpreted as demonstrating unconscious semantic activation. One of the regression method's assumptions that deserves close attention is the relative sensitivity assumption that was adapted from Reingold and Merikle (1988): the assumption that the direct measure must be at least as sensitive as the indirect measure to conscious influences of the prime stimuli. Conceivably, subjects may have been able to consciously perceive the semantic category of the prime stimuli while performing the indirect measure's task but not while performing the direct measure's task. For example, subjects might have approached the XGvWd task with the strategy of attending to physical information of the primes in order to identify Xs or Gs (rather than trying to read words). This strategy could have reduced conscious perception of semantic prime information during the direct-measure task compared to the strategy required by the indirect-measure task of attending to the meaning of the target stimuli.

Experiment 4 therefore employed a second direct measure, one for which the relative sensitivity assumption seemed unarguably acceptable. The second direct-measure task was a direct semantic-classification (DSC hereinafter) task, for which subjects were instructed to classify masked stimuli into one of the two semantic categories that had been previously used for classifying target stimuli on the indirect measure. Consequently, both the new direct measure and the indirect measure equally obliged subjects to attend to and discriminate meaning of stimuli: the former for masked stimuli in the prime position and the latter for target stimuli. Because the masked stimuli were attended in the new direct measure and ignored in the indirect measure, it seems very plausible that the direct measure should be at least as sensitive as the indirect measure to conscious effects of the meanings of those stimuli, as required by the relative sensitivity assumption. Consequently, significant intercepts in the regression analysis with the new direct measure should call unambiguously for the conclusion that priming was mediated by unconscious cognition.

Method

Subjects

Subjects were 54 undergraduate students at the University of Washington (Seattle) who volunteered to participate in the experiment in exchange for extra credit in an introductory psychology course. All were self-described as being fluent in English and having normal or corrected-to-normal vision. Two subjects were unable to complete the experiment in the scheduled 1-hr time slot, and 1 subject did not finish because of eye fatigue. Data from these 3 subjects were excluded from the analyses.

Materials and Design

In Experiment 4, the materials and design for the indirect measure's semantic-classification task were identical to those of the previous experiments. The XGvWd task (direct measure) of Experiments 1–3 was also used without changes as one of two direct measures. The new feature of Experiment 4 was the addition of the DSC task as a second direct measure.

Direct Semantic-Classification Task

The DSC task was similar to the XGvWd task in that subjects attended to the stimulus that appeared between the forward and backward mask and ignored the following distractor word. As with the XGvWd task and indirect measures, the stimulus sequence for each trial of the DSC task began with a string of 13 uppercase consonants serving as a forward mask that remained on the screen for 150 ms. The forward mask was immediately replaced by an uppercase word or name that served as the critical stimulus for this task. Critical stimuli were randomly selected from the set of words (names) used on indirect measures so that pleasant or unpleasant words (male or female names) appeared equally often for evaluative (gender) classifications. At the end of its fixed duration (17 or 33 ms), the critical stimulus was immediately replaced by a backward mask that lasted 17 ms. The backward mask was followed by a blank as needed to complete the 67-ms SOA to onset of a distractor word displayed in lowercase. Distractor words were randomly selected on each trial from the set of words (or names) used on indirect measures, with the constraint that the same stimulus never appeared as both critical stimulus and distractor on the same trial.

Subjects were instructed to ignore the clearly visible distractor words and to classify the preceding masked word as either pleasant or unpleasant in meaning (if their indirect-measure task had been

evaluative classification), or as either a male or female name (if their indirect-measure task had been gender classification). Subjects classified the stimulus as unpleasant (or male) by pressing the left response key and pleasant (or female) by pressing the right response key. As was true for the XGvWd task, no response window was used in this task.

Procedure

Each subject performed three tasks: the indirect semantic-classification task that provided the indirect (priming) measure and the two tasks (XGvWd and DSC) that provided direct measures of prime perceptibility. SOA was 67 ms for all trials, and prime durations varied randomly between 17 and 33 ms from trial to trial within each block.

After the usual practice sequence, the indirect measure's semantic-classification task was performed for six blocks of 50 test trials. The second and third tasks were the XGvWd and DSC tasks, four blocks of 50 trials each, with the order of the tasks counterbalanced across subjects. These tasks produced 135 trials for each prime duration for the indirect measure (recall that the first 5 trials of each indirect-measure block were not analyzed) and 100 trials per prime duration for each of the two direct measures.

Results

Regression Results

Priming effects were again obtained at both prime durations of 17 ms, mean $d' = 0.12$, $t(50) = 3.60$, $p = .001$, and 33 ms, mean $d' = 0.32$, $t(50) = 7.94$, $p = 10^{-10}$. As shown in Figure 5, regression analyses using the XGvWd task as the direct measure yielded significantly positive intercepts for both 17-ms, $a = 0.11$, $t(49) = 3.27$, $p = .002$, and 33-ms, $a = 0.26$, $t(49) = 6.60$, $p = 10^{-8}$, prime durations. A positive slope was obtained weakly at the 17-ms prime duration, $b = 0.15$, $t(49) = 1.33$, $p = .19$, and more strongly at 33 ms, $b = 0.23$, $t(49) = 3.35$, $p = .002$. Regression analyses using the DSC task as the direct measure, shown in Figure 6, yielded very similar findings: Statistically significant intercepts were obtained for both the 17-ms, $a = 0.12$, $t(49) = 3.59$, $p = .0008$, and 33-ms, $a = 0.31$, $t(49) = 7.88$, $p = 10^{-10}$, prime durations. (These intercepts from regressions with the DSC task were slightly larger than those obtained from the corresponding analyses with the XGvWd task.) Weakly positive slopes were obtained for both the 17-ms, $b = 0.09$, $t(49) = 0.65$, $p = .52$, and 33-ms, $b = 0.08$, $t(49) = 0.57$, $p = .57$, prime durations.

In the DSC task, it was possible that some subjects may have mistaken their task as being to classify the visible targets rather than the masked primes. Analyses were therefore conducted to determine the extent to which subjects' classifications corresponded to the semantic category of the targets. Four subjects showed responding that agreed with the target category on at least 75% of trials from both 17- and 33-ms prime-duration conditions. When these 4 subjects were dropped from regression analyses, intercepts (.12 for 17-ms primes and .32 for 33-ms primes) were almost identical to those of the full data set, although slope magnitudes (.11 for 17-ms primes and .18 for 33-ms primes) were slightly, but nonsignificantly, increased.

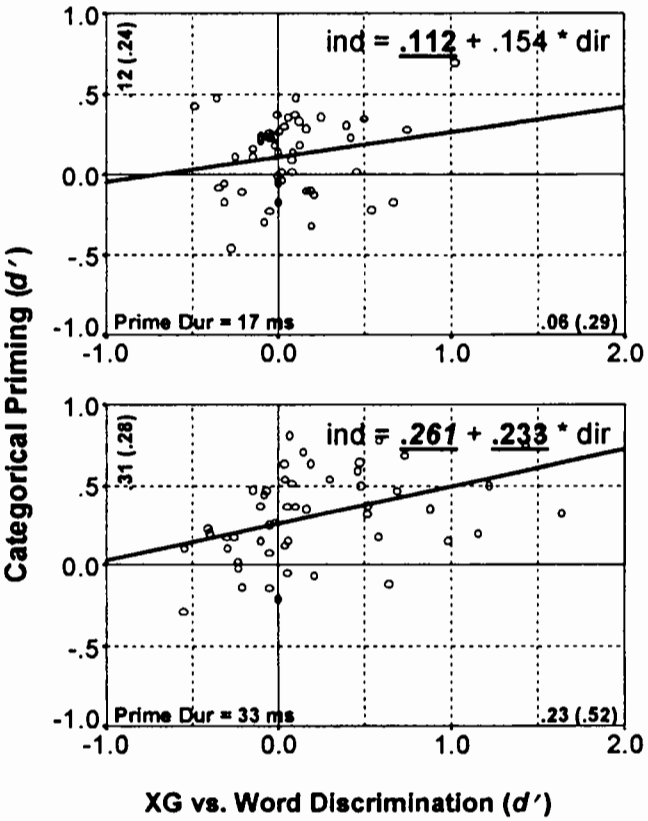

Figure 5. Scatterplots and linear-regression functions separately for the 17-ms (top) and 33-ms (bottom) prime-duration (Prime Dur) conditions of Experiment 4 ($N = 51$), using word versus XG-string discrimination as the direct measure (dir) on the abscissa and categorical priming as the indirect measure (ind) on the ordinate. Data points are individual subjects. Evaluative- and gender-classification conditions are represented indistinguishably in the plots. The direct measure (dir) of stimulus perceptibility is represented on the abscissa, and the indirect measure (ind) of semantic priming is on the ordinate. Means (and standard deviations in parentheses) for direct and indirect measures are shown in the lower right and upper left corners of the plots, respectively. Equations for best fitting linear-regression functions appear at the top of each scatterplot. In these equations, statistically significant intercept and slope parameters are printed in underlined ($p < .05$); bold and underlined ($p < .005$); or bold, underlined, and italic ($p < .0005$). The asterisk (*) indicates multiplication. XG = XG-string stimuli that were constructed from the evaluative and gender stimuli by replacing each character in those words and names alternately with an X or a G (e.g., the replacement for EVIL could be GXGX or XGXG).

Multiple Regression Analyses

Data from Experiment 4 were further analyzed using a multiple regression method in which indirect effects were simultaneously regressed onto both of the direct measures. To the extent that the two direct measures were sensitive to conscious perception of different kinds of stimulus information, the inclusion of both direct measures as predictors in the regression analysis would allow intercept effects to be

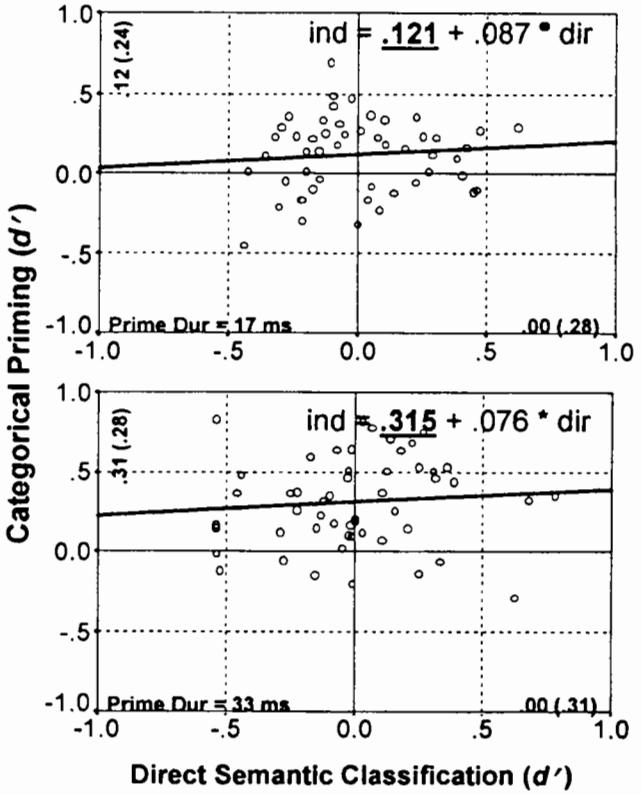

Figure 6. Scatterplots and linear-regression functions separately for the 17-ms (top) and 33-ms (bottom) prime-duration (Prime Dur) conditions of Experiment 4 (N = 51), using direct semantic classification (of primes) as the direct measure (dir) on the abscissa and categorical priming (of responses to targets) as the indirect measure (ind) on the ordinate. Data points are individual subjects. Evaluative- and gender-classification conditions are represented indistinguishably in the plots. The direct measure (dir) of stimulus perceptibility is represented on the abscissa, and the indirect measure (ind) of semantic priming is on the ordinate. Means (and standard deviations in parentheses) for direct and indirect measures are shown in the lower right and upper left corners of the plots, respectively. Equations for best fitting linear-regression functions appear at the top of each scatterplot. In these equations, statistically significant intercept and slope parameters are printed in underlined ($p < .05$); bold and underlined ($p < .005$); or bold, underlined, and italic ($p < .0005$). The asterisk (*) indicates multiplication. XG = XG-string stimuli that were constructed from the evaluative and gender stimuli by replacing each character in those words and names alternately with an X or a G (e.g., the replacement for EVIL could be GXGX or XGXG).

more confidently interpreted as corresponding to an absence of direct effects. Furthermore, even if the two direct measures were sensitive to the same stimulus information, the use of both measures would increase overall predictor reliability and heighten the precision of the intercept estimate, allowing for a more powerful test. For 17-ms primes, the multiple regression analysis yielded a significantly positive intercept, a = .11, $t(48)$ = 3.26, p = .002, along with weak positive slopes for both direct measures (XGvWd task: b = 0.16, $t(48)$ = 1.37, p = .18; DSC task: b = 0.10, $t(48)$ = 0.74, p = .46). For 33-ms primes, the regression analysis also yielded a strongly positive intercept effect, a = 0.26, $t(48)$ = 6.50, $p = 10^{-8}$, a significantly positive slope for the XGvWd task, b = 0.23, $t(48)$ = 3.26, p = .002, and a weak positive slope for the DSC task, b = 0.02, $t(48)$ = 0.19, p = .85.

Direct-Measure Performance

A comparison of the two direct measures showed that at the prime duration of 17 ms, performance on the XGvWd task (mean d' = 0.06) was slightly better than on the DSC task, mean d' = .00; D = .06, $t(50)$ = 1.16, p = .25. For 33-ms primes, performance on the XGvWd task (mean d' = 0.23) was again greater than on the DSC task (mean d' = .00; D = .23, $t(50)$ = 2.86, p = .006). These findings suggest that physical properties of the prime stimuli were more readily perceptible than semantic properties. Although Reingold and Merikle's (1988) analysis would suggest that the DSC task's direct measure was more appropriate for the regression method because of its greater superficial similarity to the indirect measure, these findings justified the additional use of the XGvWd task by showing that it was sensitive to the perception of stimulus information that was not registered by the DSC task.[5]

Analyses of Combined Data From Experiments 2–4

Figures 7 and 8 show regression analyses of evaluative- and gender-classification data for 17-ms and 33-ms prime durations, combining data from procedurally similar conditions in Experiments 2–4.[6] Experiment 1 was omitted from this aggregation because of its use of a response-window procedure (fixed 250-ms window center) that was substantially different from that of the last three experiments (longer, adaptive window center). Also omitted were the two conditions of Experiment 3 that used SOAs shorter than 67 ms (the 17/17 and 33/17 conditions). Justification for the aggregations in Figures 7 and 8 was sought from meta-analytic tests for homogeneity of effect sizes (see Rosenthal, 1993). These tests showed that the intercept effects obtained in Experiments 2–4 were not significantly heterogeneous for evaluative classifications with 17-ms primes, $\chi^2(2, N = 131) = 1.93$, p = .38, or 33-ms primes, $\chi^2(2, N = 131) =$

[5] A reviewer of a previous draft of this article commented that fatigue or boredom may have reduced sensitivity of direct measures, which were always collected after indirect measures. However, analyses of Experiment 4's direct measures (the order of which was counterbalanced) showed no significant performance advantage for the first-performed task over the second, $F(1, 50)$ = 1.98, p = .17. Furthermore, collecting direct measures before indirect measures is generally inadvisable because direct measures alert subjects to the presence and temporal location of the prime stimuli, increasing the likelihood that subjects might inappropriately attempt to consciously process the primes while performing the indirect-measure task.

[6] Regression analyses of the combined evaluative- and gender-priming data sets from Experiments 2–4 were reported in Greenwald, Draine, and Abrams (1996).

1.24, $p = .54$, or for gender classifications with 17-ms primes, $\chi^2(2, N = 131) = 2.40$, $p = .30$, or 33-ms primes, $\chi^2(2, N = 131) = 0.58$, $p = .75$.

Linear-Regression Results

Although statistically significant intercept effects were obtained in only two of the four experiments for 17-ms primes, the more powerful aggregated analysis of evaluative classifications in Experiments 2–4 yielded a significant positive intercept, $a = 0.08$, $t(64) = 2.80$, $p = .007$, and a weakly positive slope, $b = 0.15$, $t(64) = 1.63$, $p = .11$. Similarly, the aggregated analysis of gender classifications with 17-ms primes in Experiments 2–4 yielded a regression function with a significant positive intercept, $a = 0.08$, $t(62) = 2.70$, $p = .009$, and a weakly positive slope, $b = 0.03$, $t(62) = 0.33$, $p = .74$. For 33-ms primes, the aggregated regression analysis involving evaluative classifications yielded a positive and highly significant intercept, $a = 0.27$, $t(64) = 7.84$, $p = 10^{-10}$. The slope was also positive and statistically significant, $b = 0.13$, $t(64) = 2.12$, $p = .04$. The aggregated regression analysis involving gender classifications yielded a positive and highly significant intercept, $a = 0.34$, $t(62) = 9.09$, $p = 10^{-12}$, and a weakly positive slope, $b = 0.08$, $t(62) = 1.19$, $p = .24$.

Within both the 17-ms and 33-ms prime-duration conditions, parameter estimates from the aggregated regression analyses involving evaluative classifications were remarkably similar in magnitude to those involving gender classifications. These similarities justified the strategy of combin-

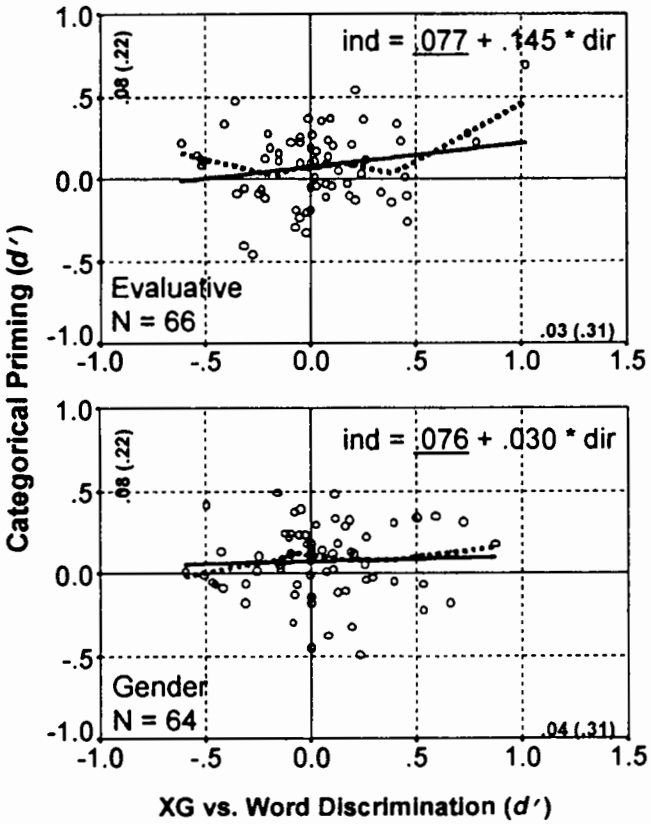

Figure 7. Scatterplots and linear-regression functions for aggregated data of the 17-ms prime-duration conditions in Experiments 2–4, separately for evaluative- (top) and gender- (bottom) classification tasks, using word versus XG-string discrimination as the direct measure (dir) on the abscissa and categorical priming as the indirect measure (ind) on the ordinate. Nonparametric lowess smoother curves are superimposed on the scatterplots to facilitate identification of deviations from linearity. Data points are individual subjects. Evaluative- and gender-classification conditions are represented indistinguishably in the plots. The direct measure (dir) of stimulus perceptibility is represented on the abscissa, and the indirect measure (ind) of semantic priming is on the ordinate. Means (and standard deviations in parentheses) for direct and indirect measures are shown in the lower right and upper left corners of the plots, respectively. Equations for best fitting linear-regression functions appear at the top of each scatterplot. In these equations, statistically significant intercept and slope parameters are printed in underlined ($p < .05$); bold and underlined ($p < .005$); or bold, underlined, and italic ($p < .0005$). The asterisk (*) indicates multiplication. XG = XG-string stimuli that were constructed from the evaluative and gender stimuli by replacing each character in those words and names alternately with an X or a G (e.g., the replacement for EVIL could be GXGX or XGXG).

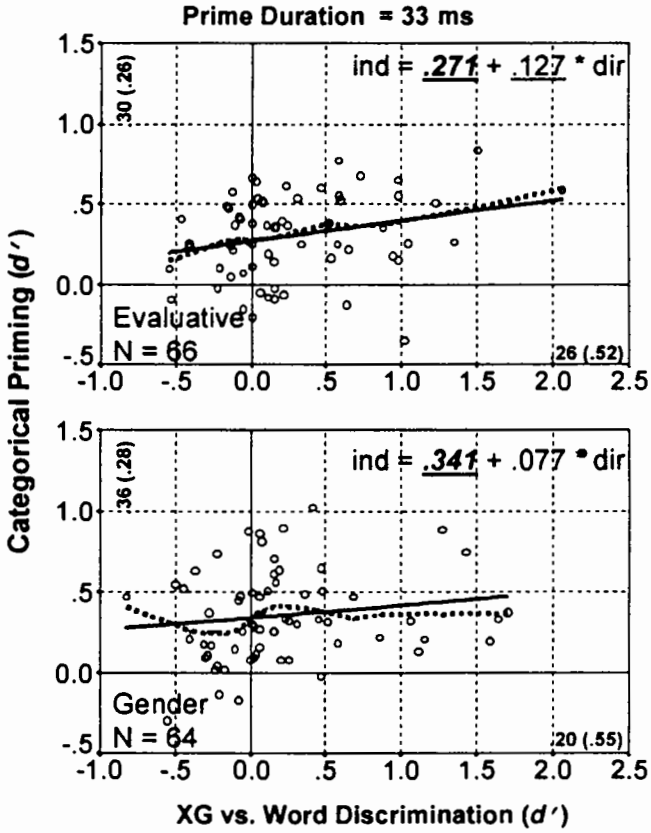

Figure 8. Scatterplots and linear-regression functions for aggregated data of the 33-ms prime-duration conditions in Experiments 2–4, separately for evaluative (top) and gender (bottom) classification tasks, using word versus XG-string discrimination as the direct measure (dir) on the abscissa and categorical priming as the indirect measure (ind) on the ordinate. The asterisk (*) indicates multiplication.

ing data from both tasks in the prior presentation of results for the individual experiments.

Nonlinear-Regression Results

Nonparametric lowess smoother curves (Chambers, Cleveland, Kleiner, & Tukey, 1983) were examined in order to assess whether the relations between direct and indirect measures seriously violated the assumption of linearity. The lowess curves were strings of best fitting linear-regression functions estimated separately at each data point by fitting only the nearest 65% of all points along the abscissa. The lowess curves involving 17-ms primes (shown in Figure 7) gave some indication that indirect and direct measures may not have been strictly linearly related as assumed by the linear-regression analyses. Nevertheless, the lowess curves for both evaluative and gender classifications crossed the *y*-axis at nearly the same point as did the corresponding linear-regression function. Thus, violations of the linearity assumption did not appear to threaten the conclusion of positive intercept effects. For 33-ms primes, the lowess curves for both evaluative and gender classifications (shown in Figure 8) appeared generally consistent with the assumption that direct and indirect measures were linearly related and again supported the conclusion of positive intercept effects. For both 17-ms and 33-ms prime durations, lowess curves that were more locally weighted also indicated positive intercepts of similar magnitude.

Retroactive Semantic Priming?

A number of investigators have reported evidence of retroactive semantic priming (Briand, den Heyer, & Dannenbring, 1988; Dark, 1988; VanVoorhis & Dark, 1995): the facilitation of perceptual and semantic judgments of masked stimuli by subsequently presented, semantically related, and visible stimuli. In the present experiments, it was possible that the visible distractors appearing at the end of direct-measure trials may have retroactively primed the subliminally presented critical stimuli, producing higher direct effects on congruent trials (prime and target from the same semantic category) than on incongruent trials (prime and target from different semantic categories).

On half of the trials of the XGvWd task, prime stimuli were words that were semantically congruent or incongruent with the to-be-ignored target stimuli. Data from this subset of trials from Experiments 2–4 were analyzed to determine whether prime–target relations influenced the percentage of correct classifications of word stimuli. The results, shown in Table 1, indicated that no differences in the percentage of correct word classifications were found between congruent and incongruent trials. Thus, the XGvWd task provided no evidence for retroactive semantic priming..

On the DSC task, retroactive semantic priming would similarly be expected to produce a higher percentage of correct semantic classifications of prime words on trials with congruent rather than incongruent distractors. As shown in Table 1, direct semantic classifications were indeed more accurate on congruent than incongruent trials. This result, however, can be explained in terms of cognitive phenomena other than retroactive priming. For example, if on a given trial subjects did not consciously perceive any semantic information of the prime, they would have had to respond by guessing. (The finding of d' values near zero on the DSC task suggests that subjects may have had to guess quite frequently.) The semantic content of the target stimuli may have influenced guesses if (a) subjects adopted the strategy of guessing with the target's semantic category as a default or (b) if guesses were indirectly primed by the semantic content of the target. The present data set offers no clear means of choosing between these possible explanations.

Table 1
Retroactive Semantic Priming on Direct Measures

Stimulus duration	Congruent % correct	Incongruent % correct	*N*	*p*
XG–word discrimination				
17 ms	60.58	61.45	173	.21
33 ms	58.30	58.79	173	.40
50 ms	61.69	59.39	80	.54
Direct semantic classification				
17 ms	56.79	43.20	51	<.001
33 ms	57.04	42.93	51	<.001

Measurement Error on Direct Measures

A general assumption of regression analyses is that the predictor variable is free of measurement error. Violations of this assumption cause the absolute magnitude of the estimated regression slope to be smaller than that of the true slope, that is, positive slopes tend to be underestimated, and negative slopes tend to be overestimated (Cohen & Cohen, 1983). To the extent that (a) the true slope of the regression analysis is greater than zero and (b) the overall mean score on the predictor variable is greater than zero, this "flattening" of the slope by predictor error causes the estimated intercept to be greater than the true intercept (see Figure 7 of Greenwald & Draine, 1997). In the context of the present research, an artificially inflated intercept might lead to the mistaken conclusion that evidence for unconscious cognition (i.e., the indirect-without-direct-effect data pattern) had been obtained.

As indicated by the occurrence of negative d' values, the direct measures used in the present research clearly contained measurement error and therefore necessarily violated the measurement-error-free predictor assumption of regression analyses. However, examination of the results of Experiments 1–4 shows that intercept effects were obtained in conditions with mean performance on the direct measure near zero (see Figure 4 [top and bottom], Figure 6 [top], and Figure 7 [top and bottom]) and in conditions with regression slopes that were nearly flat (see Figure 1 [bottom], Figure 2 [middle and bottom], Figure 3 [bottom], Figure 4 [top], and Figure 6 [top and bottom]) or negative (see Figure 1 [middle], Figure 3 [top], and Figure 4 [bottom]). These intercept effects appeared in conditions that either lacked the

properties necessary to inflate intercept estimates or had properties (i.e., a negative slope) that would have actually deflated the intercept estimates from their true values. Therefore, the intercept effects cannot reasonably be dismissed as artifacts of measurement error in the direct measure.

General Discussion

Replicable Subliminal Priming

Until the present research, reproducible evidence for subliminal priming in the form of the indirect-without-direct-effect data pattern had eluded researchers, despite concerted effort since the early 1980s. In each of the four present experiments, the indirect-without-direct-effect data pattern appeared in the form of a statistically significant positive intercept effect. The key methodological device for producing this effect was use of a response-window procedure, which obliged target-classification response judgments to be made more rapidly than with previously standard procedures for assessing priming effects with latency measures. The response-window procedure allowed priming effects to be focused on the performance dimension of accuracy, rather than being distributed (and diluted) across both speed and accuracy measures. By concentrating the priming effect on the accuracy dimension, sensitivity to priming was increased.

Response-Window Procedure Focuses Priming Effects on Performance Accuracy

Data showing strong priming effects in accuracy measures have already been considered in detail (see Figures 1–8). If the response-window procedure effectively focused priming effects onto the performance accuracy dimension, then response latencies should have been affected relatively little by congruence versus incongruence of the priming stimulus. Table 2 shows mean latencies for congruent and incongruent trials separately for prime-duration conditions of 17, 33, and 50 ms, aggregated across Experiments 2–4. (As for the regression analyses reported earlier, data from Experiment 1 and from conditions of Experiment 3 with SOAs shorter than 67 ms were excluded from the aggregated analyses.) No evidence of priming appeared in latency analyses for 17- and 33-ms primes: Mean latencies did not significantly differ between congruent and incongruent trials. In the 50-ms prime-duration condition, however, latencies did show a small (11 ms) but statistically significant priming effect. Despite the evidence for some leakage of priming effects onto the latency measure with 50-ms primes, the response-window procedure was apparently quite effective in concentrating priming effects in the accuracy component of responding.

Response-Window Procedure and Magnitude of Subliminal Semantic Activation

The response-window procedure not only focused priming effects onto the accuracy dimension. It also produced considerably larger priming effects than have been obtained in previous research on subliminal semantic activation. Regression analyses of data aggregated across Experiments 2–4 yielded intercept effects of $d' = 0.08$ for 17-ms primes, $d' = 0.31$ for 33-ms primes, and $d' = 0.44$ for 50-ms primes. Expressed as **d** (computed as the intercept divided by the standard deviation of the indirect measure), these intercepts corresponded to effect sizes of **d** = 0.34, **d** = 1.14, and **d** = 1.38, respectively. The intercept effects for the 33- and 50-ms prime-duration conditions easily surpassed the value (**d** = 0.80) associated with the conventional designation as a large effect size. Further, these intercept effects are an order of magnitude larger than the statistically significant intercept effects previously reported by Greenwald et al. (1995) and Greenwald and Draine (1997) from studies that did not use the response-window procedure.

Table 2
Target Classification Latencies (ms) for Incongruent and Congruent Prime–Target Pairs by Prime Duration From Experiments 2–4

Presentation condition	Incongruent		Congruent		Incong–cong with 95% C.I.
	M	*SD*	*M*	*SD*	
17 ms primes	426	66	427	69	-1 ± 3.8^a
33 ms primes	429	71	426	66	3 ± 4.0^a
50 ms primes	439	77	428	69	11 ± 8.3^b

Note. Incong = incongruent; cong = congruent; C.I. = confidence interval.
[a] $N = 130$. [b] $N = 37$.

Response-Window Procedure and Time-Course of Subliminal Semantic Activation

It seems likely that the response-window procedure increased priming effects by constraining responses to occur during a period of heightened susceptibility to influence by primes. Consider that on each trial, primes and targets served as two orthogonal (by experimental design) sources of activation along the task-relevant semantic dimensions of evaluation or gender. Because primes always preceded targets in the presentation sequence, the onset and subsequent cumulation of prime-triggered activation necessarily preceded the onset and cumulation of target-triggered activation. By obliging subjects to respond even more rapidly to the target than do standard reaction time instructions, the response-window procedure may have elicited responding at a time that was (a) enough delayed after prime onset for prime-triggered activation to have accumulated to a relatively high level, yet (b) not so delayed that this activation would have decayed to a low level, and also (c) too soon after target onset for target-triggered activation to have reached a high enough level to yield highly accurate responding.

If the response-window procedure indeed produced a confluence of the three properties just described, it would have created conditions in which the influence of the prime stimulus would be much greater than in standard priming

procedures. By contrast, standard priming procedures leave the subject free to withhold a response long enough to permit both greater decay of prime-triggered activation and greater accumulation of target-triggered activation. It can be seen in Table 2 that latencies under the response-window procedure of the present research averaged about 430 ms, which can be compared with typical values between 550 and 650 ms in otherwise similar priming experimental procedures. This comparison suggests that the addition of little more than 100 ms to response latencies may provide enough time for the relative strengths of prime-initiated and target-initiated semantic activation to change dramatically. In support of the speculation that activation induced by masked primes can peak and diminish rapidly, Greenwald et al. (1996)—using a response-window procedure with a priming task similar to that of the present research—found that priming effects were much smaller for prime–target SOAs exceeding 100 ms than for SOAs between 67 ms and 100 ms.

Other Possibly Important Procedural Factors

In addition to using the response window, the present research used several other procedures that might have contributed to finding strong subliminal priming. Perhaps most important was the use of prime–target SOAs that were short in comparison with those used in many previous subliminal priming studies. The present SOAs, which were uniformly 67 ms or less, approximated the range of values at which Greenwald et al. (1996) were able to obtain strong subliminal priming effects. These SOAs were also much shorter than the values of 250 to 2,000 ms that have been typical of most previous investigations of automatic priming (e.g., Balota, 1983; Neely, 1977).

Another procedure that distinguished the present research from many previous investigations of subliminal priming was its use of semantic classification as the assigned task for target stimuli. It is possible that the semantic-classification task is better suited to producing subliminal semantic priming than are other tasks such as the lexical decision task used by Dagenbach, Carr, and Wilhemsen (1989) and Brown and Hagoort (1993), the color-naming Stroop task used by Cheesman and Merikle (1984), or the position-discrimination priming task used by Greenwald et al. (1995).

Still another procedural factor that differentiated the present research from previous, less successful efforts to produce subliminal priming was its use of both forward and backward masking of prime stimuli. This masking technique (informally referred to as "shutter masking" or "sandwich masking") permitted use of longer prime durations than could be used in previous subliminal priming experiments that have relied mostly on backward masking alone. The present finding that priming magnitude increased with prime duration gives some support to the suggestion that the present relatively long prime durations may have contributed to the ease with which subliminal priming was observed in this research.

Theoretical Implications of Intercept Effects

The intercept effects repeatedly obtained in Experiments 1–4 are compatible with two theories of unconscious cognition: (a) association—unconscious cognition exists, but only in conjunction with conscious cognition; and (b) dissociation—unconscious cognition exists independently of conscious cognition. According to the association view, the intercept effects reflect unconscious semantic activation that occurred in combination with at least some amount of conscious processing of the semantic category of the prime. In contrast, the more controversial dissociation view maintains that semantic activation occurred in conditions that prevented any conscious processing of the primes' semantic categories.

If it is assumed that direct measures were sensitive to all consciously mediated processing of the semantic category of the primes (i.e., an exhaustiveness assumption), then the significant intercept effects would support dissociation over association. (See Greenwald et al., 1995, p. 37, for discussion of relevance of the exhaustiveness assumption to the dissociation interpretation.) Both the XGvWd and the DSC tasks should have been sensitive in principle to any conscious processing of the semantic category of the primes because any such processing would have provided enough information to guide a correct response on either direct-measure task. Although the assumption of exhaustive sensitivity to conscious semantic processing is plausibly valid for the direct measures, the present experiments provided no empirical confirmation of this assumption.

Conclusion

Subliminal semantic activation (SSA) has heretofore been widely treated as a dubious phenomenon. Skeptics have been inclined to interpret claimed findings of SSA as artifacts resulting from procedurally flawed research. Even those sympathetic to the phenomenon have been obliged to acknowledge that it has been empirically recalcitrant. The present research takes SSA research beyond this stage of questioning validity by establishing it as a robustly replicable phenomenon. Empirically, this research demonstrates methods that, by sharply magnifying priming effects, can greatly expedite the testing of theories about mechanisms of priming. The central ingredient of these methods is the device of constraining response latencies to values 100 ms or more faster than those at which subjects can maintain high accuracy levels. Theoretically, these findings suggest that subliminal priming has previously been difficult to observe because semantic activation induced by visually masked priming dissipates very rapidly.

References

Attneave, F. (1959). *Applications of information theory to psychology.* New York: Holt, Rinehart & Winston.

Avant, L. L., & Thieman, A. A. (1985). On visual access to letter case and lexical/semantic information. *Memory & Cognition, 13,* 393–404.

Balota, D. A. (1983). Automatic semantic activation and episodic memory encoding. *Journal of Verbal Learning and Verbal Behavior, 22,* 88–104.

Bellezza, F. S., Greenwald, A. G., & Banaji, M. R. (1986). Words high and low in pleasantness as rated by male and female college students. *Behavior Research Methods, Instruments, and Computers, 18,* 299–303.

Bonanno, G. A., & Stillings, N. A. (1986). Preference, familiarity, and recognition after repeated brief exposures to random geometric shapes. *American Journal of Psychology, 99,* 403–415.

Bornstein, R. F., & D'Agostino, P. A. (1992). Stimulus recognition and the mere exposure effect. *Journal of Personality and Social Psychology, 63,* 545–552.

Briand, K., den Heyer, K., & Dannenbring, G. L. (1988). Retroactive semantic priming in a lexical decision task. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 40A,* 341–359.

Brown, C., & Hagoort, P. (1993). The processing nature of the N400: Evidence from masked priming. *Journal of Cognitive Neuroscience, 5,* 34–44.

Chambers, J. M., Cleveland, W. S., Kleiner, B., & Tukey, P. A. (1983). *Graphical methods for data analysis.* Boston: Duxbury Press.

Cheesman, J., & Merikle, P. M. (1984). Priming with and without awareness. *Perception & Psychophysics, 36,* 387–395.

Cheesman, J., & Merikle, P. M. (1986). Distinguishing conscious from unconscious perceptual processes. *Canadian Journal of Psychology, 40,* 343–367.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Dagenbach, D., Carr, T. H., & Wilhelmsen, A. (1989). Task-induced strategies and near-threshold priming: Conscious influences on unconscious perception. *Journal of Memory and Language, 28,* 412–443.

Dark, V. J. (1988). Semantic priming, prime reportability, and retroactive priming are interdependent. *Memory & Cognition, 16,* 299–308.

Dixon, N. F. (1971). *Subliminal perception: The nature of a controversy.* London: McGraw-Hill.

Dosher, B. A., & Rosedale, G. (1991). Judgments of semantic and episodic relatedness: Common time-course and failure of segregation. *Journal of Memory and Language, 30,* 125–160.

Doyle, J. R., & Leach, C. (1988). Word superiority in signal detection: Barely a glimpse yet reading nonetheless. *Cognitive Psychology, 20,* 283–318.

Dunkling, L., & Gosling, W. (1984). *The facts on file dictionary of first names.* New York: Facts on File Publications.

Eriksen, C. W. (1960). Discrimination and learning without awareness: A methodological survey and evaluation. *Psychological Review, 67,* 279–300.

Fodor, J. A. (1983). *The modularity of mind.* Cambridge, MA: MIT Press.

Fowler, C. A., Wolford, G., Slade, R., & Tassinary, L. (1981). Lexical access with and without awareness. *Journal of Experimental Psychology: General, 110,* 341–362.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics.* New York: Wiley.

Greenwald, A. G. (1992). New look 3: Unconscious cognition reclaimed. *American Psychologist, 47,* 766–779.

Greenwald, A. G., & Draine, S. C. (1997). Do subliminal stimuli enter the mind unnoticed? Tests with a new method. In J. Cohen & J. Schooler (Eds.), *Scientific approaches to consciousness* (pp. 83–108). Mahwah, NJ: Erlbaum.

Greenwald, A. G., Draine, S. C., & Abrams, R. H. (1996, September). Three cognitive markers of unconscious semantic activation. *Science, 273,* 1699–1702.

Greenwald, A. G., & Klinger, M. R. (1990). Visual masking and unconscious processing: Differences between backward and simultaneous masking? *Memory & Cognition, 18,* 430–435.

Greenwald, A. G., Klinger, M. R., & Liu, T. J. (1989). Unconscious processing of dichoptically masked words. *Memory & Cognition, 17,* 35–47.

Greenwald, A. G., Klinger, M. R., & Schuh, E. S. (1995). Activation by marginally perceptible ("subliminal") stimuli: Dissociation of unconscious from conscious cognition. *Journal of Experimental Psychology: General, 124,* 22–42.

Grice, G. R., Nullmeyer, R., & Spiker, V. A. (1977). Application of variable criterion theory to choice reaction time. *Perception & Psychophysics, 22,* 431–449.

Groeger, J. A. (1988). Qualitatively different effects of undetected and unidentified auditory primes. *Quarterly Journal of Experimental Psychology, 40A,* 323–329.

Hirshman, E., & Durante, R. (1992). Prime identification and semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 255–265.

Holender, D. (1986). Semantic activation without conscious identification in dichotic listening, parafoveal vision, and visual masking: A survey and appraisal. *Behavioral and Brain Sciences, 9,* 1–23.

Jacoby, L. L., Lindsay, D. S., & Toth, J. P. (1992). Unconscious influences revealed: Attention, awareness, and control. *American Psychologist, 47,* 802–809.

Kostandov, E. A. (1985). Neurophysiological mechanisms of "unaccountable" emotions. In J. T. Spence & C. E. Izard (Eds.), *Motivation, emotion, and personality* (pp. 175–193). Amsterdam: North-Holland.

Kunst-Wilson, W. R., & Zajonc, R. B. (1980, February). Affective discrimination of stimuli that cannot be recognized. *Science, 207,* 557–558.

Mandler, G., Nakamura, Y., & Van Zandt, B. J. S. (1987). Nonspecific effects of exposure on stimuli that cannot be recognized. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 646–648.

Marcel, A. J. (1980). Conscious and preconscious recognition of polysemous words: Locating the selective effects of prior verbal context. In R. S. Nickerson (Ed.), *Attention and performance, VIII* (pp. 435–457). Hillsdale, NJ: Erlbaum.

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

Merikle, P. M. (1984).Towards a definition of awareness [Abstract]. *Bulletin of the Psychonomic Society, 22,* 449–450.

Merikle, P. M., & Reingold, E. M. (1991). Comparing direct (explicit) and indirect (implicit) measures to study unconscious memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 224–233.

Meyer, D., & Schvaneveldt, R. (1971). Facilitation in recognizing pairs of words: Evidence of dependence between retrieval operations. *Journal of Experimental Psychology, 90,* 227–234.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General, 106,* 226–254.

Ollman, R. (1977). Choice reaction time and the problem of distinguishing task effects from strategy effects. In S. Dornic (Ed.), *Attention and performance VI* (pp. 99–113). Hillsdale, NJ: Erlbaum.

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola Symposium* (pp. 55–85). Hillsdale, NJ: Erlbaum

Reed, A. V. (1973, August). Speed–accuracy trade-off in recognition memory. *Science, 181,* 574–576.

Reingold, E. M., & Merikle, P. M. (1988). Using direct and indirect measures to study perception without awareness. *Perception & Psychophysics, 44,* 563–575.

Rosenthal, R. (1993). Cumulating evidence. In G. Keren & C. Lewis (Eds.), *A handbook for data analysis in the behavioral sciences: Methodological issues* (pp. 519–559). Hillsdale, NJ: Erlbaum.

Seamon, J. G., Marsh, R. L., & Brody, N. (1984). Critical importance of exposure duration for affective discrimination of stimuli that are not recognized. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 465–469.

Shevrin, H. (1988). Unconscious conflict: A convergent psychodynamic and electrophysiological approach. In M. J. Horowitz (Ed.), *Psychodynamics and cognition* (pp. 117–167). Chicago: University of Chicago Press.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643–662.

VanVoorhis, G. A., & Dark, V. J. (1995). Semantic matching, response mode, and response mapping as contributors to retroactive and proactive priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 913–932.

Wickelgren, W. A. (1977). Speed–accuracy tradeoff and information processing dynamics. *Acta Psychologica, 41,* 67–85.

Appendix

Stimuli

Evaluative stimuli		Gender stimuli	
Unpleasant	Pleasant	Male	Female
evil	honor	mike	sarah
cancer	lucky	david	kate
sickness	diamond	jason	becky
disaster	loyal	jeff	laura
poverty	freedom	john	julie
vomit	rainbow	mark	jill
bomb	love	kevin	eve
rotten	honest	bob	jean
abuse	peace	bill	joan
murder	heaven	derek	sally
assault	pleasure	brad	april
slime	family	karl	vicki
divorce	diploma	tom	jenny
poison	kiss	matt	pam
kill	cheer	adam	lisa
death	health	nick	amy
hatred	friend	eric	emily
scum	caress	steve	tara
accident	sunset	brian	holly
jail	happy	joe	ann
stink	miracle	scott	jane
torture	sunrise	harry	mary
crash	paradise	paul	diana
filth	vacation	bart	alice
pollute	treasure	alan	tammy

Received March 7, 1996
Revision received July 31, 1996
Accepted January 10, 1997 ■