Psychological Bulletin
2009, Vol. 135, No. 3, 452–477

0033-2909/09/$12.00 DOI: 10.1037/a0015329

Mechanisms of Masked Priming: A Meta-Analysis

Eva Van den Bussche, Wim Van den Noortgate, and Bert Reynvoet
University of Leuven

The extent to which unconscious information can influence behavior has been a topic of considerable debate throughout the history of psychology. A frequently used method for studying subliminal processing is the masked priming paradigm. The authors focused on studies in which this paradigm was used. Their aim was twofold: first, to assess the magnitude of subliminal priming across the literature and to determine whether subliminal primes are processed semantically, and second, to examine potential moderators of priming effects. The authors found significant priming in their analyses, indicating that unconsciously presented information can influence behavior. Furthermore, priming was observed under circumstances in which a nonsemantic interpretation could not fully explain the effects, suggesting that subliminally presented information can be processed semantically. Nonetheless, the nonsemantic processing of primes is enhanced and priming effects are boosted when the experimental context allows the formation of automatic stimulus–response mappings. This quantitative review also revealed several moderators that influence the strength of priming.

Keywords: masked priming, subliminal priming, semantic priming, meta-analysis

Can unconsciously presented information influence behavior? This is one of the most controversial questions in the history of psychology and remains an intriguing and strongly debated question today (e.g., Merikle & Daneman, 1998; Sidis, 1898). Numerous researchers have investigated the topic by presenting stimuli below the *limen* or the threshold of conscious perception and assessing whether these subliminal stimuli influence behavior. One of the earliest descriptions of subliminal priming was reported in 1898 by Sidis. He showed participants cards with letters or numbers on them at such a distance that they claimed to be unable to see what was on the cards; nonetheless, they performed above chance when guessing the cards' identities. This and other early studies, however, have been met with considerable skepticism. Still, the issue of unconscious perception continues to intrigue us and to receive considerable empirical attention (see Kouider & Dehaene, 2007, for an extensive review).

For almost 3 decades, the masked priming paradigm has been used to study the impact of subliminal information on behavior. The paradigm involves assessing the influence of a masked stimulus on the processing of a subsequent target. For instance, Marcel (1983) found that target words were processed faster when they were preceded by a semantically related prime word (e.g., *cat–dog*) than by an unrelated word (e.g., *book–dog*), even when the primes were rendered subliminal by masking them and presenting them for only a very short duration. Although Marcel's results were first described as "startling" and "counterintuitive" (Fowler, Wolford, Slade, & Tassinary, 1981, p. 341), successful replications accumulated throughout the years (e.g., Balota, 1983; Forster & Davis, 1984; Fowler et al., 1981; Greenwald, Klinger, & Liu, 1989). But as the pile of supporting evidence grew, so too did the skepticism. Holender (1986) reviewed the use of masked priming and concluded that the findings were problematic in a variety of ways, including lack of reliability and poor assessment of whether stimuli were actually presented subliminally. The identification of serious methodological flaws caused great doubt as to the existence of subliminal processing.

By the mid-1990s, the development of new and stronger paradigms prompted a renewed interest in the topic. In 1998, Dehaene and colleagues asked participants to classify numbers between 1 and 9 as smaller or larger than 5. Two types of trials were presented: *congruent* trials (e.g., 1–3) in which primes and targets evoked the same response and *incongruent* trials (e.g., 1–8) in which primes and targets evoked different responses. They found that responses were faster on congruent than on incongruent trials, a phenomenon called the *response congruency effect*. Moreover, their study was the first to use brain-imaging techniques, which showed that subliminal primes elicited neural activity in the motor cortex. According to Dehaene et al., this proved that participants also unconsciously applied the task instructions to the subliminal primes when they performed the semantic categorization task. The authors concluded that subliminal primes are processed in a series of stages, including *semantic processing*. The fact that subliminal primes facilitated the subsequent categorization of targets belong-

Eva Van den Bussche and Bert Reynvoet, Laboratory of Experimental Psychology, Department of Psychology, University of Leuven, Leuven, Belgium, and Subfaculty of Psychology and Educational Sciences, University of Leuven–Campus Kortrijk, Kortrijk, Belgium; Wim Van den Noortgate, Subfaculty of Psychology and Educational Sciences, University of Leuven–Campus Kortrijk.

We wish to thank all researchers who contributed to this meta-analysis by providing us with the statistical information needed to compute effect sizes. We especially want to thank Kenneth Forster, Andrea Kiesel, Sachiko Kinoshita, and Carsten Pohl for sending us unpublished data and Tom Verguts for reading a previous version of this article and offering his useful comments.

Correspondence concerning this article should be addressed to Eva Van den Bussche, Faculty of Psychology and Educational Sciences, University of Leuven–Campus Kortrijk, E. Sabbelaan 53, 8500 Kortrijk, Belgium. E-mail: Eva.Vandenbussche@kuleuven-kortrijk.be

ing to the same semantic category suggested that the primes were unconsciously categorized and processed semantically. This and other investigations, all using sound methodological approaches (see also Draine & Greenwald, 1998; Greenwald, Draine, & Abrams, 1996), removed the stigma that had surrounded research on subliminal priming. Nowadays, the existence of subliminal perception is largely acknowledged. However, the debate has progressed beyond existence claims. Ever since Dehaene et al. (1998) interpreted their findings as proof of semantic processing of subliminal information, the depth of subliminal priming rather than its existence has been the issue of interest.

In 2001, Damian shed new light on Dehaene and colleagues' findings. Dehaene et al. had used prime stimuli that also appeared as target stimuli. Damian investigated the possibility that the use of stimuli as both primes and targets affected the results. His findings confirmed this possibility. The response congruency effect disappeared when primes and targets were completely different stimuli, and thus the primes never elicited a response. According to Damian, the lack of a congruency effect when primes were never presented as targets indicated that participants learned to associate stimuli directly with appropriate responses, creating automatized stimulus–response (S–R) mappings that bypassed semantic access. A prime can automatically activate the corresponding response without first accessing semantics when it is also presented as a target (e.g., Dehaene et al., 1998). Such an S–R mapping cannot be established, however, when participants never respond to a prime (e.g., Damian, 2001), which explains the absence of a response congruency effect (see Abrams & Greenwald, 2000, for a similar claim). This provided an alternative, nonsemantic explanation of the response congruency effect. Although S–R mappings can explain the response congruency effect when primes are also used as targets, it cannot explain response congruency effects observed in other studies in which primes and targets were completely different (e.g., Dell'Acqua & Grainger, 1999; Klauer, Eder, Greenwald & Abrams, 2007; Naccache & Dehaene, 2001; Reynvoet, Gevers, & Caessens, 2005). These findings are difficult to explain without assuming that subliminal primes are semantically processed.

Hitherto, the debate as to whether subliminal priming reflects genuine semantic processing of the subliminal information or the formation of automatic S–R mappings remains undecided. Still, it is important to be able to distinguish *response priming* and true *semantic priming*. Response priming is based on S–R mappings that are established during the experiment, whereas true semantic priming is based on preexisting semantic associations between primes and targets (Kiefer, 2007). Furthermore, differential underlying neural substrates can be identified for response and semantic priming. As Kiefer (2007) argued, it is important to assess whether the unconscious processes underlying these two distinct forms of priming are governed by the same top-down influences (e.g., task sets, intentions).

Divergent research results have also given rise to the supposition that several factors moderate subliminal priming effects. Kouider and Dehaene (2007) recently provided an excellent historical overview of the literature on masked priming; however, the field still lacks a quantitative review of the available data, one in which the overall effect size is assessed and the influence of potential moderators is investigated. Our aim in the current study was to combine published and unpublished data through meta-analytic techniques to examine several unresolved issues. First, we examined whether the literature provides clear evidence in favor of semantic processing of subliminal information by estimating and testing the overall effect size of subliminal priming. Second, we investigated the influence of potential moderators of subliminal priming. A unique aspect of meta-analysis is that it yields statistical tests of moderating effects. This can shed light on the conditions under which subliminal processing is or is not observed. A meta-analysis, such as this one, can make important contributions to the field of subliminal priming research: (a) It can help clarify when and to what depth subliminal primes are processed, providing data that have implications for competing theories; (b)it can guide future research, identifying important gaps in the literature and suggesting potential moderators and interactions; and (c) it can provide important methodological information, identifying factors that are crucial in designing subliminal priming experiments.

Our meta-analysis focused exclusively on subliminal priming studies in which the prime stimulus is made "invisible" by *masking*. Visual masking is a method used to reduce the visibility of a stimulus by presenting another visual stimulus in close temporal and spatial proximity. In a typical subliminal priming experiment, a trial starts with the presentation of a forward mask, followed by the very brief presentation of a prime stimulus. Then, a backward mask is shown, followed by the target stimulus. In another variant (e.g., Forster & Davis, 1984), the prime is immediately replaced by the target. Under such conditions, participants have difficulty reporting the identity of the prime stimulus.

In the following section, we provide a detailed description of the moderators that we included in our meta-analysis. This list is not meant to be exhaustive, since we could include only those moderators that were frequently reported in the literature.

Tasks

Three standard tasks are used to examine subliminal priming. In a *semantic categorization task,* participants are asked to decide whether a visible target belongs to one semantic category or another. For example, participants may be asked to categorize numbers as smaller or larger than 5. A subliminal prime precedes the target and belongs either to the same semantic category as the target (i.e., congruent trial, e.g., 1–3) or to another semantic category (i.e., incongruent trial, e.g., 1–8). The priming effect, also called the response congruency effect, is manifested as faster or more accurate responses on congruent than on incongruent trials. In a *lexical decision task,* participants receive letter strings and are asked to make word/nonword judgments. Subliminal primes precede the targets and are semantically related (e.g., *doctor–nurse*) or unrelated (e.g., *butter–nurse*) to the target words. The priming effect is manifested as faster or more accurate responses to semantically related prime–target pairs than to unrelated pairs. A *naming task* is similar to the lexical decision task except that the targets are all words, and participants are asked to name the targets aloud. Priming is defined in the same way as in the lexical decision task.

A consistent claim in the literature is that the size of the priming effect depends on the nature of the task. For example, Lucas (2000) found that priming effects in naming were smaller than those in lexical decision. Similarly, priming in semantic categorization appears to be stronger than priming in lexical decision (e.g., Grainger & Frenck-Mestre, 1998). The latter observation can be explained by the fact that categorization requires access to semantic information, whereas naming and lexical decision do not. Furthermore, priming in semantic categorization is biased by response congruency (e.g., Forster, 2004). Primes and targets belong to the same semantic category and require the same response on congruent trials, whereas primes and targets belong to different categories and require different responses on incongruent trials. If participants apply task instructions to the primes as well as to the targets (Dehaene et al., 1998), then priming may not originate at the level of the category but may originate at the response level, involving *response priming* rather than semantic priming. This confound is usually absent in naming and lexical decision in which both related and unrelated primes belong to the same semantic category and require a similar response. Thus, these latter tasks may reveal true *semantic priming*. It should be noted, however, that response congruency can be manipulated in naming and lexical decision tasks, for example, by presenting words and nonwords as both primes and targets (e.g., Klinger, Burton, & Pitts, 2000). These kinds of response congruency manipulations in lexical decision and naming were not considered in the current meta-analysis.

Prime Novelty

Primes can also appear as targets (i.e., repeated primes; e.g., Dehaene et al., 1998), or primes can never be used as targets (i.e., novel primes; e.g., Damian, 2001). The completely opposite nature of results obtained with repeated primes and novel primes suggests that prime novelty may moderate priming effects such that priming is more likely to occur with repeated primes than with novel primes.

Category Size

Some researchers have found priming only when stimuli were members of small categories (e.g., body parts, numbers, months). When stimuli came from large categories, such as animals, priming was not observed (Forster, 2004; Forster, Mohan, & Hector, 2003). This suggests that priming may decrease as a function of category size.

Target Set Size, Target Repetitions, and Number of Trials

Kiesel, Kunde, Pohl, and Hoffmann (2006) did obtain priming using a large category but only when they presented a large number of targets. When a limited number of targets were presented, no priming was found. These results are consistent with findings reported by Abrams (2008) who also found no priming for novel primes when the target set size was small. In addition to target set size, the number of times that targets are repeated may also moderate subliminal priming. Damian (2001), for example, claimed that the formation of S–R links was facilitated by presenting the same targets repeatedly. Finally, the number of trials in an experiment may influence priming, since lengthy tasks can induce fatigue, diminish concentration, and possibly decrease priming. Alternatively, increasing the number of trials may reduce noise in the data and yield more robust priming effects.

Prime and Target Format and the Correspondence Between Them

In subliminal priming experiments, the primes and targets can take various forms: words, digits (e.g., 1, 2), number words (e.g., one, two), pictures, Chinese symbols, or letters. An extensive body of naming research documents differences in the processing of word, picture, and symbol stimuli (see Glaser, 1992, for a review). For visually presented words, naming can occur without semantic mediation, whereas pictures and symbols (digits, Chinese characters) cannot be named unless their meanings are activated (e.g., Perfetti & Tan, 1998; Tan, Hoosain, & Peng, 1995; Tan, Hoosain, & Siok, 1996; Theios & Amrhein, 1989; Ziegler, Ferrand, Jacobs, Rey, & Grainger, 2000). This discrepancy implies that the prime and target format might moderate priming. We can also ask whether the correspondence between prime and target format influences priming. Kunde, Kiesel, and Hoffmann (2003) showed that number word primes did not elicit priming when participants received only digit targets, suggesting that the correspondence between prime and target format might moderate subliminal priming.

Prime Duration and Stimulus Onset Asynchrony

One of the defining characteristics of subliminal priming is that the prime is presented below the threshold for conscious perception. One means to assure this is to present primes for very short durations. Several studies have found that priming increases with prime duration (e.g., Holcomb, Reder, Misra & Grainger, 2005; Klauer et al., 2007). Furthermore, response priming increases monotonically as a function of stimulus onset asynchrony (SOA, i.e., the interval between the onset of the prime and the onset of the target) (Vorberg, Mattler, Heinecke, Schmidt, & Schwarzbach, 2004). In contrast, semantic priming decreases at long SOAs (e.g., Greenwald et al., 1996; Kiefer & Spitzer, 2000).

Masking

A second means of ensuring subliminal presentation of a prime is masking. Several masking techniques are common. Only *backward* masking is of interest in the current study since the goal of backward masking is to eliminate the visual image of the prime by replacing it with a new image, whereas the purpose of *forward* masking is to alert participants that a new trial is beginning. One backward masking method is *pattern* masking, in which a visual pattern follows the prime and is then replaced by the target. A pattern mask can be a series of symbols (e.g., #### or %#%#), letters, or scrambled patterns. In a second masking method, which we will call backward *target* masking, the prime is followed immediately by presentation of the target. The nature of the masking procedure may be a factor that influences subliminal priming. Klauer et al. (2007) reported minimal priming in their study and suggested that severe masking conditions may have reduced the amount of priming (see also Van Opstal, Reynvoet, & Verguts, 2005b).

Visibility Measures

The subliminal presentation of primes is the definitive feature of subliminal priming; thus, studies should provide some measure of whether this criterion was met. If no measure of prime visibility was reported, we have no way of knowing whether the primes were indeed presented subliminally. If studies reported strong priming but failed to provide a prime visibility measure, then we might suspect that the primes were presented above threshold (e.g., Greenwald, Abrams, Naccache, & Dehaene, 2003; Holcomb et al., 2005). Thus, the presence or absence of a prime visibility measure may moderate priming effects.

In general, measures of prime awareness can be subdivided into two classes (Merikle & Reingold, 1992): measures assessing a *subjective* threshold, in which conscious awareness is indexed by participants' self reports (e.g., asking whether participants were aware of the primes) and measures assessing an *objective* threshold, in which conscious awareness is indexed by a measure of the participants' discriminative abilities, such as a forced-choice absent–present decision or a categorization task (e.g., asking participants to categorize the subliminal primes instead of the targets). An objective measure provides a stricter criterion for conscious awareness than does a subjective measure, leading to more conservative evaluations of prime visibility (e.g., Cheesman & Merikle, 1986). The adequacy and reliability of these measures are strongly debated (see, for example, Merikle & Reingold, 1992); nonetheless, we can assume that priming should diminish when stringent (i.e., objective) methods of assessing prime visibility are used.

Some studies have reported a direct measure of prime visibility (d' measure). The d' measure is a sensitivity measure based on signal detection theory (Greenwald et al., 2003). After responding to the targets in the usual priming task, participants receive the same prime–target stimuli but then respond to the primes. Mean d' values are positively related to prime visibility. Unawareness of the primes is assumed when d' values do not differ significantly from 0. Thus, d' measures may moderate priming effects such that stronger priming is associated with greater visibility as indexed by d'.

Study Features

We also included sample size and population as potential moderators. If publication or reporting bias is present in the subliminal priming literature, we expected that smaller samples would lead to stronger priming effects. In addition, priming may be larger in some populations than others.

Method

Literature Search

Five procedures were used to retrieve published and unpublished studies (Lipsey & Wilson, 2001). First, we conducted database searches of PsycINFO, Web of Science, PubMed, and ScienceDirect (which also covers Dissertation Abstracts International). We used the search string (*semantic or associative) and (priming or prime) and (masked or subliminal or unconscious or automatic*). We used language and publication date as additional search parameters: Studies were considered for inclusion only when they were published in English, French, Dutch, or German between January 1983 (year of Marcel's seminal publication) and December 2006. Second, we searched the tables of content in journals that commonly publish articles on this topic (e.g., *Advances in Cognitive Psychology*, *Cognition*, *Journal of Experimental Psychology: Human Perception and Performance,* and *Consciousness and Cognition*). Third, we examined the reference sections of relevant literature reviews for additional citations. Fourth, we checked the reference sections of all potentially qualifying articles for citations. Finally, we contacted established subliminal priming researchers and requested relevant published and unpublished studies.

With this search strategy, we identified 749 studies (see Figure 1). Most studies contained multiple conditions (e.g., one study often consisted of several experiments, and each experiment often contained several manipulations of potential moderators). We defined a *study* as a published or unpublished collection of one or more experiments. A *condition* is nested within a study and involves manipulations that are relevant to the aims of our meta-analysis. Please note that a condition is not necessarily equivalent to an experiment: Often one experiment contained multiple relevant conditions. We included these conditions separately in our analyses. Therefore, we used the term *condition* rather than *experiment*.

Inclusion and Exclusion Criteria

We used the following criteria to select studies and conditions within studies for inclusion in the meta-analysis. Please note that an entire study was excluded only when all conditions within it were excluded on the basis of these criteria.

1. The relation between the prime and target was of a direct semantic nature (for example, *cat–dog*) in the visual domain. Thus, we excluded conditions investigating phonological priming, morphological priming, orthographic priming, negative priming, repetition priming, cross-language priming, stem completion, mediated priming, action priming, and auditory priming.

2. The primes were presented subliminally. Thus, conditions were excluded when the prime was presented for 100 ms or more, when a task was performed on the prime, when participants were explicitly aware of the primes, or when a prime was insufficiently masked because a long blank (and no backward mask) occurred between prime and target.

3. Our meta-analysis focused solely on studies that used the most common experimental tasks in priming: semantic categorization, lexical decision, and naming. Conditions were excluded when other tasks were used (i.e., rapid serial visual presentation, Stroop or flanker tasks, or double tasks) or when no experimental priming task was used (e.g., questionnaire studies, reviews, or theoretical articles).

4. Conditions were included only when they involved a standard priming procedure, in which participants were instructed to respond as fast as possible to the target that was preceded by a subliminal prime and response times were the dependent variable. Conditions were excluded when the procedure deviated from the standard (i.e., cueing procedures, mood induction procedures, or procedures including a study phase) or when a different dependent variable was measured (e.g., response window procedures).

5. We also limited the analysis to centrally presented single word or symbol primes, in which primes and targets appeared at the same location to minimize the influence of spatial attention. Conditions were excluded when primes and targets appeared at different locations and

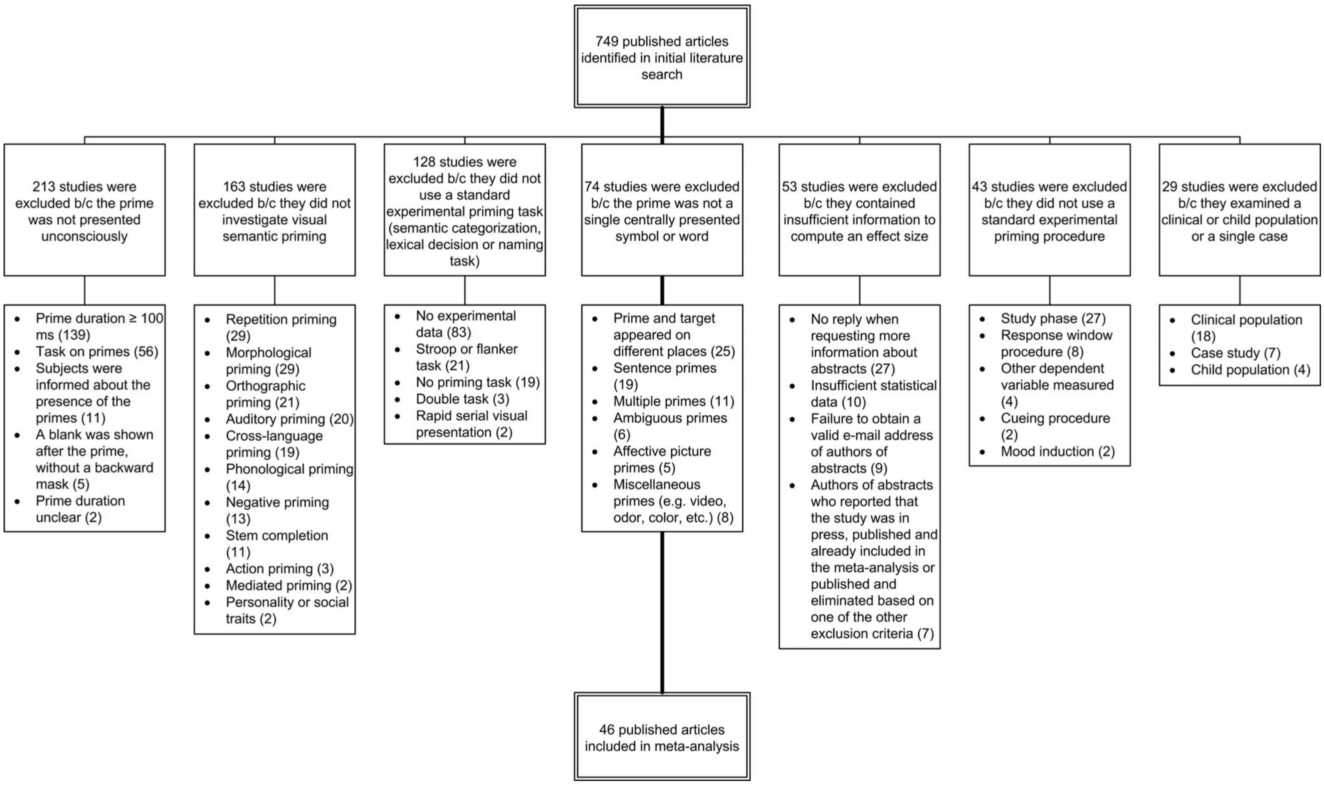

Figure 1. Flowchart illustrating the number of published articles omitted based on the seven inclusion criteria. b/c = because.

when conditions involved sentence primes, color primes, ambiguous primes, multiple primes, video primes, odor primes, music primes, context primes, or two-letter strings. Conditions were excluded when affective picture primes were used because affective pictures, but not affective words, may be processed via a subcortical route (LeDoux, 1996).

6. We focused exclusively on conditions with healthy adult samples of more than 1 participant. Conditions were excluded when they involved children, case studies, or clinical populations without a control group.

7. Finally, conditions were included only when the studies reported sufficient statistical information to compute an effect size. If insufficient information was reported, the corresponding author was contacted to obtain the necessary data. The condition was excluded from the meta-analysis when two attempts to obtain the information were unsuccessful. Furthermore, we contacted all authors of potentially relevant conference abstracts when the information could not be obtained from one of the articles that we found in our search. Conditions were excluded when authors did not reply, when we were unable to obtain a valid e-mail address, or when the authors informed us that the abstract was now an article in press, had been published as a full article that was already included in the meta-analysis, or had been published as a full article that was eliminated due to any of the other exclusion criteria.

An overview of the excluded studies as a function of the seven criteria is shown in Figure 1. We note that this figure shows only the most important reason for excluding a study (since several studies were excluded for multiple reasons). The application of these criteria resulted in the selection of 46 studies published between 1983 and 2006 and 8 unpublished studies.

Coding Procedure and Reliability

We used a standard coding system to rate each condition. Appendix A lists the moderators that were coded for each condition and explains how each moderator was operationalized. The appendix also includes descriptive statistics for each moderator. Some moderator categories were grouped to limit the number of levels (as described in the appendix).

We obtained intercoder agreement for each moderator by having two independent reviewers code the moderators for a randomly selected 30% of the studies. We calculated the intraclass correlation coefficient for continuous variables and kappa coefficients for categorical variables. The intraclass correlation coefficients ranged from .997 (for target set size) to 1.0 (for all other continuous variables). The kappa coefficients ranged from .811 (for the prime visibility measure) to 1.0 (for all other categorical variables). The intercoder agreement was high because the coding was very straightforward. Disagreements were resolved by discussion; the final coding reflects the consensus between the two raters.

Meta-Analytic Procedures

Only conditions that met all inclusion criteria were included in the meta-analysis. At least one condition met all inclusion criteria in 54 studies. In total, these studies yielded 156 separate conditions. We made a strict distinction between semantic categorization tasks and lexical decision and naming tasks because of possible differences in how priming originated and because conditions in which lexical decision or naming was used had very different features than did those in which semantic categorization was used. This can be seen in Appendixes B and C. For example, almost all conditions using lexical decision and naming tasks involved word primes and targets, novel primes, and large stimulus categories, whereas a more balanced pattern was seen in conditions with semantic categorization tasks. We divided the 156 conditions into two groups: one group contained semantic categorization tasks (23 studies comprising 88 separate conditions); the other group contained lexical decision and naming tasks (32 studies comprising 68 separate conditions). Note that one study contained both semantic categorization and lexical decision and naming; thus, it was included in both groups. Two separate meta-analyses were performed on the two data sets, always using the same procedure. Appendixes B and C provide an overview and description of the moderators included in the semantic categorization analyses (Appendix B) and lexical decision and naming analyses (Appendix C).

All studies used a single-group repeated measures design, in which the same participants received multiple treatment conditions (Morris & DeShon, 2002). Each participant's reaction time (RT, in milliseconds) was measured on related or congruent (e.g., *cat—dog*) and on unrelated or incongruent (e.g., *cat—pot*) trials. A widespread method of calculating effect sizes for independent-groups designs (in which the outcome is measured at a single point in time and is compared across independent groups that receive different treatments; see Hedges, 1981; 1982) cannot be applied in these designs. In a single-group repeated measures design, a difference score can be calculated for each participant as the mean RT on unrelated/incongruent trials minus the mean RT on related/congruent trials. We used the mean of these differences divided by the standard deviation of the differences to express the priming effect in each condition (Gibbons, Hedeker, & Davis, 1993) as follows:

$$\hat{\theta} = \frac{M_D}{SD_D},$$

in which $\hat{\theta}$ is the estimated effect size for condition j in study k,[1] M_D is the difference between the mean RTs in unrelated/incongruent and related/congruent trials (unrelated/incongruent − related/congruent), and SD_D is the sample deviation of the differences.

Most studies reported the means of the observed priming effects; however, the standard deviations were often unavailable. Therefore, we always estimated the effect sizes from reported test statistics. We were able to obtain repeated measures t test or F test statistics for all conditions. These test statistics can be transformed into a repeated measures effect size with the following conversion formula (Rosenthal, 1991):

$$\hat{\theta} = \frac{t}{\sqrt{n}}$$

or

$$\hat{\theta} = \sqrt{\frac{F}{n}},$$

in which n is the number of participants. The square root of the F value does not indicate the direction of the difference; thus, we specified the positive or negative direction based on the pattern of means. A positive effect size was interpreted as a positive priming effect (i.e., faster responses to related/congruent than to unrelated/incongruent prime–target pairs).

We also estimated the sampling error (*SE*) for each condition before conducting the meta-analyses. The inverse of the estimated sampling variance of the observed effect sizes was used to weight the effect sizes. For single-group repeated measures designs, the following variance formula has been proposed (Morris & DeShon, 2002):

$$SE^2 = \left(\frac{1}{n}\right)\left(\frac{n-1}{n-3}\right)(1 + n\hat{\theta}^2) - \frac{\hat{\theta}^2}{[c(df)]^2},$$

in which n is the number of participants and the bias function $c(df)$ is approximated by (Hedges, 1982):

$$c(df) = 1 - \frac{3}{4(n-2)}.$$

Appendixes B and C provide an overview of the effect sizes and corresponding sampling errors for the semantic categorization (Appendix B) and lexical decision and naming (Appendix C) data sets separately.

We conducted the meta-analyses after estimating the effect size ($\hat{\theta}_{jk}$) and its corresponding sampling error SE_{jk} for each condition j in study k. The nesting of participants within conditions and conditions within studies yields three potential sources of variance. Two of these are present in the typical meta-analysis: (a) *sampling variance* (i.e., differences between observed effect sizes and population effects) and (b) *between-study variance* (i.e., systematic differences between the population effect sizes from different studies). A third source of variance is (c) *between-condition/within-study variance* (i.e., systematic differences between the effect sizes from different conditions within the same study). We used the multilevel meta-analysis approach to account for the three sources of variance (see Van den Noortgate & Onghena, 2003). Raudenbush and Bryk (1985) showed that a meta-analysis is a special case of multilevel analysis, except that aggregated instead of raw data are used. Following the multilevel research tradition, we started with a random-effects model (REM), in which studies are a random sample from a population of studies rather than direct replications of each other. The REM without moderators is shown below:

$$\hat{\theta}_{jk} = \beta_0 + V_k + U_{jk} + e_{jk},$$

[1] The standard symbol used to describe the sample estimator of the effect size is d. However, to avoid all confusion with d', the objective measure of prime visibility mentioned in this study, we will use the symbol $\hat{\theta}$ to denote the sample statistic.

in which $\hat{\theta}_{jk}$ is the observed effect size for condition j in study k; β_0 is the overall mean effect size across all conditions and studies; V_k refers to the random deviation of the effect in study k from the overall effect; U_{jk} refers to the deviation of the effect for condition j in study k from the mean effect in study k; and e_{jk} is the residual due to sampling fluctuation, indicating the deviation of the observed effect size from the population effect size for condition j in study k. All three error terms, V_k, U_{jk}, and e_{jk}, are assumed to be independently and normally distributed with 0 mean. Note that the sampling variance for each condition was estimated before the meta-analyses were conducted (see previous text). Thus, only β_0, the overall mean effect size, σ^2_V, the between-study variance component, and σ^2_U, the within-study variance component, were estimated in the meta-analysis.

We can extend this REM by including moderators. Restricted maximum likelihood estimation was used to estimate the parameters, as implemented in the mixed procedure from SAS (Littell, Milliken, Stroup, Wolfinger, & Schabenberger, 2006). The estimates from the individual studies were automatically weighted by the reciprocal of the variance of the observed effect sizes; in our case, this was the sum of the sampling variance, the between-condition/within-study variance, and the between-study variance. In this approach, more precise observed effect sizes have greater impact on the results. Furthermore, this kind of variance weighting accounts for both sample size and study design (Hedges & Olkin, 1985).

We investigated whether the variance between the observed effect sizes was larger than what would be expected on the basis of sampling variance alone to determine whether the effect sizes were homogeneous (i.e., if they could be treated as estimates of a common effect size). If the effect sizes are heterogeneous, then moderators of the effects are likely to be present. We used a likelihood ratio test, comparing models with and without the between-study variance component (σ^2_V). The difference between the two deviance scores, defined as -2 times the log likelihood, follows a 50–50 mixture of the chi square distributions for 0 and 1 degrees of freedom (Raudenbush & Bryk, 2002; Snijders & Bosker, 1999), making it possible to determine whether significant variance is present at the between-study level. The same procedure can be used to test the significance of the between-condition/within-study variance (σ^2_U). If these tests indicate significant between-study and/or between-condition/within-study variance, moderators of the effect sizes are likely. The multilevel meta-analysis model can include multiple moderator variables, without assuming that all heterogeneity between studies and between conditions can be explained by the included moderators (Van den Noortgate & Onghena, 2003).

The same sample was often used across multiple conditions within a study, so that some samples contributed more than one effect size to the analyses. Specifically, 46 of the 156 conditions had a sample that was used in another condition. This means that a substantial portion of the effect sizes within a study were not independent. Rosenthal (1991) noted that the inclusion of multiple effect sizes from the same sample treats dependent results as independent, weighting each study according to the number of effect sizes it produces. The outcome of the meta-analysis can be biased if there is anything unusual or unrepresentative about those studies that contribute more than one effect size. To take this into account, we should ideally conduct a multivariate analysis in which we also include the *sampling covariance*. The data necessary to calculate the sampling covariance (i.e., estimates of the covariances between the effect sizes within a study), however, were neither reported in the studies that we included nor listed elsewhere in the literature. The number of conditions that included a sample used in multiple conditions was substantial; thus, we decided not to restrict each effect size to a single sample to avoid an extensive loss of information. However, we conducted several sensitivity analyses to study the impact of the dependency on the outcome patterns (Greenhouse & Iyengar, 1994). For both the semantic categorization and the lexical decision and naming meta-analyses, all analyses were conducted twice including only one randomly selected effect size per sample. This allowed us to investigate how ignoring the dependency between the effect sizes influenced the results.

Another important issue is the possibility of publication bias. Significant results are more likely to be published than null results (Begg, 1994). If the meta-analysis is limited to published studies, the true mean effect size may be overestimated. One method for avoiding publication bias is to include as many unpublished studies as possible. In the current meta-analyses, eight unpublished studies (i.e., 15%) comprising 37 conditions (i.e., 24%) were included: five unpublished studies (i.e., 22%) comprising 29 conditions (i.e., 33%) in the semantic categorization meta-analysis and three unpublished studies (i.e., 9%) comprising 8 conditions (i.e., 12%) in the lexical decision and naming meta-analysis. We did not consider manuscripts under revision or articles in press to ensure that, as far as possible, the unpublished data were not soon-to-be-published data (and thus likely to report mainly significant results). The most important means of identifying publication bias is sample size: small studies produce highly variable effect size estimates; aberrant values that occur by chance are much farther from the true mean effect size for small studies than for large studies. Thus, effect sizes are likely to be more positive from small samples than from large ones if publication bias is present (Begg, 1994). This should lead to a negative correlation between sample size and effect size: larger effect sizes for small, more unreliable studies and smaller effect sizes for large, more reliable studies. We examined publication bias by constructing "funnel graphs," on which sample size is plotted as a function of effect size (Light & Pillemer, 1984). The plot should be shaped like a funnel if no bias is present, with the spout pointing up (a broad range of points for the variable small studies at the bottom and a narrow range of points for the large studies at the top). The mean effect size should be similar regardless of sample size: When a vertical line is drawn through the mean effect size, the graph should be symmetrical on either side of the line. The funnel will be skewed, however, if publication bias is present. We also used a second method for detecting publication bias; we included a moderator that coded for whether the study was published or unpublished and tested its influence on effect sizes.

SEMANTIC CATEGORIZATION

Results

Descriptive Statistics

We identified 23 studies that met all inclusion criteria and used a semantic categorization task in at least one condition. In total, 88 conditions were extracted from these 23 studies. All conditions included an academic or a nonacademic sample with a mean sample size of 20. Primes and targets were either symbols, words, or a mixture of both. Primes and targets were presented in the same format in most conditions. Primes were presented for an average of 42 ms with an average SOA of 106 ms. Participants received novel primes, repeated primes, or a mixture of both. An average of 21 targets were presented per condition; each was repeated an average of 68 times. The average number of trials in a condition was 485. The stimuli were from small categories in half of the conditions and from large categories in the other half. Most of the conditions included pattern masking, and the majority assessed prime visibility. A d' measure was reported in two thirds of the conditions. Appendix A provides descriptive statistics for the 88 semantic categorization conditions as a function of the potential moderators.

Overall Mean Effect Size and Effect Size Heterogeneity

Appendix B contains the observed effect sizes and corresponding sampling errors for the 88 conditions. Across all conditions, the mean effect size for the random effects model was 0.80 (k = 88, 95% confidence interval [CI] = 0.60−1.00). Figure 2 graphically displays all observed effect sizes ordered by size and the overall mean effect size with 95% CIs.

We conducted a likelihood ratio test comparing models with and without between-study variance that showed that significant variance (i.e., Var) was present at the between-study level, Var = 0.15, $\chi^2(1, k = 88) = 18.1, p < .0001$.[2] In addition, we found significant differences between conditions within studies, Var = 0.10, $\chi^2(1, k = 88) = 27.4, p < .0001$. Thus, moderators of the effect sizes are likely.

We assessed how much of the variance was situated at each of the three levels in the meta-analysis by using the median sampling variance (i.e., median Var) to calculate the total amount of variance (sum of the between-study variance, the within-study/between-condition variance, and the median sampling variance). This strategy was used because the sampling variances vary depending on study and condition within study. We estimated the sampling variance for each condition before conducting the meta-analyses. This meant that we could not readily determine how much of the variance was located at the third level. We solved this problem using the median sampling variance across all conditions. The median amount of sampling variance was 0.10. Thus, the total sum of the between-study variance, the within-study/between-condition variance, and the median sampling variance was 0.35 (0.15 + 0.10 + 0.10). On the basis of this, we calculated the percentage of variance situated at the between-study level (42%) and at the within-study/between-condition level (29%). According to the 75% rule devised by Hunter and Schmidt (1990), between-study or between-condition/within-study variances are substantial, and the influence of potential moderators should be examined if less than 75% of the variance is due to sampling error (or other artefacts). On the basis of this rule, we can again conclude that the effect sizes are heterogeneous, suggesting that moderators may account for the variability in effect sizes.

Regression Models With One Moderator

We first examined random effect regression models in which each of the 16 moderators (see Appendix A) was entered separately. Table 1 provides an overview of these models. Below, we provide a description of the significant moderators and then list the nonsignificant ones.

Prime format. Prime format was a significant moderator of the observed effect sizes. Pairwise comparisons indicated that the average effect size for word primes (0.51) was significantly smaller than the average effect size for symbol primes (0.91) and mixtures of symbol and word primes (1.15), $t(54.4) = 2.70$, $p = .009$, and $t(30.1) = 3.25$, $p = .003$, respectively. Nonetheless, significance tests of the regression coefficients (t tests) against the null mean showed that the average effect sizes for all three kinds of primes were significantly different from 0 (see Table 1). Prime format explained 20% of the variance (calculated across the three levels and using the median sampling variance).

Prime novelty. Prime novelty was a second significant moderator of the effect sizes. The average effect size for novel primes was smaller than the average effect size for repeated primes: 0.57 and 1.08, $t(85) = 5.02$, $p < .0001$, respectively. Nonetheless, significance tests of the regression coefficients showed that the average effect sizes for novel primes, repeated primes, and the mixture of both were all significantly different from 0 (see Table 1). Prime novelty explained 23% of the variance.

Target format. A third moderator of the effect sizes was target format. The average effect size was smaller for word targets (0.49) than for symbol targets (1.01) and mixtures of symbol and word targets (1.08), $t(35.9) = 3.56$, $p = .001$, and $t(33) = 2.95$, $p = .006$, respectively. Nonetheless, significance tests of the regression coefficients showed that the average effect sizes for all three kinds of targets were different from 0 (see Table 1). Target format explained 24% of the variance.

Category size. Category size was a strong moderator of the effect sizes. The average effect size for conditions using stimuli from a large category was smaller than the effect size for those using small categories: 0.38 and 1.09, $t(14.5) = 6.55$, $p < .0001$, respectively. Again, significance tests of the regression coefficients indicated that the average effect sizes for small and large categories were both different from 0 (see Table 1). Category size explained 39% of the variance.

The d' measure. Fifty-eight of the 87 conditions included a d' measure of prime visibility. The measure was a significant moderator of the effect sizes, with effect sizes increasing as a function of d' (i.e., higher prime visibility), $\beta = 1.04$, $F(1, 54.1) = 8.09$, $p = .006$; d' explained 20% of the variance. The positive relation of d' to the effect sizes indicates that effect sizes (and thus the observed priming effects) increased as prime visibility increased. To assess whether a significant effect size was still present when the visibility of the primes was 0, we tested whether the intercept was significant. This was indeed the case, $\beta = 0.44$, $t(14.3) = 3.73$, $p = .002$, indicating that significant priming can be expected even when prime awareness is presumably absent.

Other moderators had no significant influence on effect sizes. These included population, sample size, target set size, target repetitions, number of trials, prime–target format, prime duration, SOA, masking, whether or not prime visibility was measured, and

[2] In fact, a 50–50 mixture of the chi square distributions for 0 and 1 degrees of freedom was used, but we always simply refer to this as $\chi^2(1)$.

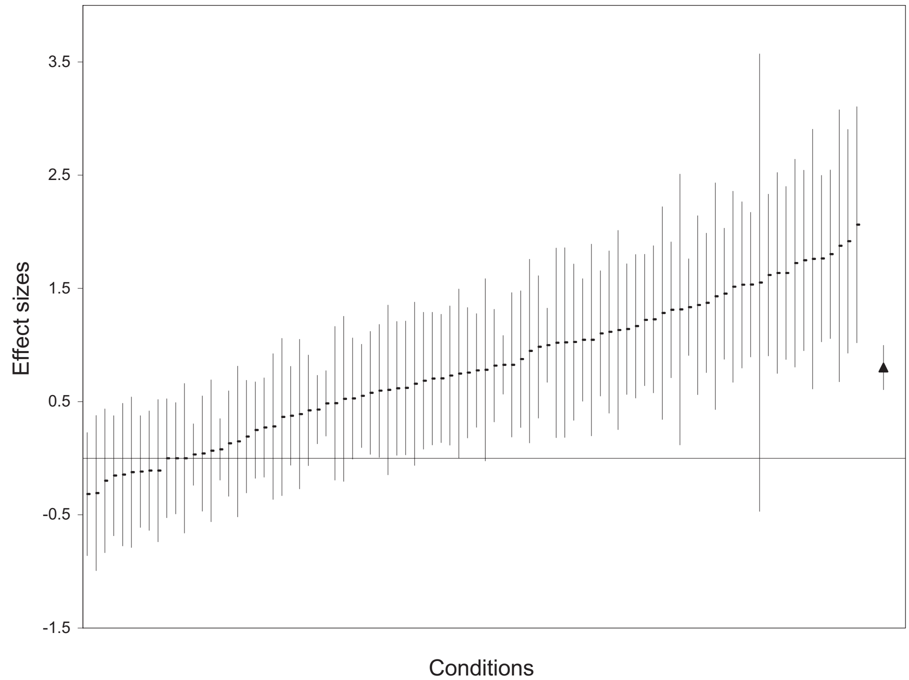

Figure 2. Observed effect sizes for the semantic categorization conditions ordered by size of the effect size and the overall mean effect size (indicated by a triangle) with their 95% confidence intervals.

the nature of the visibility measure. More details are provided in Table 1.

Regression Models With Two-Way Interactions

After identifying the five significant moderators, we examined two-way interactions among them. We added target set size because it was explicitly suggested as a potential moderator by Kiesel et al. (2006). We entered main effects and interactions between pairs of moderators in the analyses. None of the two-way interactions were significant (p values ranging from .37 to .98). No higher order interactions were investigated because some moderator categories contained only a few observations.

Multiple Regression Models

Pearson correlations among the moderators were computed to obtain an overview of the relations among them. We transformed categorical variables into dichotomous ones by eliminating the level with the least observations (this procedure was used only for these correlational analyses). This involved eliminating the category "both" for the following variables: prime format, target format, prime novelty, and the nature of the visibility measure. As a result, 64 of the 88 conditions were retained for analysis. The correlations are shown in Table 2. The table also contains the variance inflation factors (VIF). The VIF is a measure of multicollinearity; VIF values greater than 4–10 generally indicate severe multicollinearity (O'Brien, 2007). Please note that d' was not included in calculating VIF values because this would have further reduced the number of conditions to 44. The VIF values slightly exceeded the cutoff value for prime format (VIF = 6.49) and target format (VIF = 7.80). The substantial overlap between these two moderators makes sense, since most studies used the same kind of primes and targets within a condition. This was confirmed by the very large correlation between the two variables (r = .87). Prime format is an important theoretical variable in our meta-analyses; thus, we decided to omit target format from the multiple moderator analyses. After we eliminated target format, all remaining VIF values fell well below the cutoff (VIF range 1.35–3.59). The observed overlap among some variables suggests confounding; thus, care should be taken in interpreting the results from the regression models in which only one moderator was included.

Only main effects were included in the multiple regression models because we sought to enhance interpretability and because no two-way interactions were significant. We used a backward elimination strategy, starting with a model that contained all moderators and then eliminating nonsignificant ones step by step on the basis of their p value (the moderator with the highest p value was eliminated first). Because in the estimation of the model parameters conditions are only included if data are available for all moderators in the model (conditions with missing values are excluded), the moderators d' and the nature of the visibility measure were excluded because they could only be computed on a reduced number of effect sizes (k = 58 and k = 74, respectively). The R^2 for the full model (k = 88) was .45; only prime novelty and category size were significant, $F(2, 61.6) = 8.61$, $p = .0005$, and $F(1, 26.4) = 4.82$, $p = .04$, respectively. The consistent Akaike information criterion (CAIC) was used to test whether the full model performed better than an empty model containing no moderators. The CAIC is a goodness-of-fit measure that adjusts the model chi square test to penalize for model complexity and sample size. This measure can be used to compare nonhierarchical and hierarchical (nested) models. Lower values indicate better fit

Table 1
Regression Models With One Moderator for Semantic Categorization Conditions

Moderator	k	β	95% CI	F	df	p	Model R^2
N	88	−0.0009	−0.01, 0.01	0.03	1, 37.6	.87	.00
Population	88			1.61	1, 31.1	.21	.01
Academics	63	0.73****	0.50, 0.95				
Nonacademics	25	0.98****	0.64, 1.33				
Prime format	88			6.57	2, 41.4	.003	.20
Symbols	32	0.91****	0.67, 1.16				
Words	38	0.51***	0.30, 0.73				
Both	18	1.15****	0.83, 1.48				
Prime novelty	88			13.60	2, 66.2	<.0001	.23
Repeated	40	1.08****	0.88, 1.29				
Novel	44	0.57****	0.38, 0.76				
Both	4	0.71**	0.22, 1.20				
Target format	88			7.73	2, 40.4	.001	.24
Symbols	36	1.01****	0.78, 1.24				
Words	38	0.49***	0.29, 0.69				
Both	14	1.08****	0.74, 1.41				
Category size	88			42.92	1, 14.5	<.0001	.39
Small	44	1.09****	0.93, 1.24				
Large	44	0.38***	0.23, 0.52				
Set size	88	−0.002	−0.008, 0.004	0.48	1, 35.3	.50	.02
Target repetitions	88	−0.0002	−0.002, 0.002	0.03	1, 75.0	.86	.00
Trials	88	0.0001	−0.0005, 0.007	0.11	1, 47.9	.74	.00
Prime–target format	88			0.25	1, 78.7	.62	.00
Same	80	0.79****	0.58, 0.99				
Different	8	0.90***	0.45, 1.36				
Prime duration	88	0.008	−0.006, 0.02	1.40	1, 34.5	.24	.01
SOA	88	0.0004	−0.002, 0.003	0.14	1, 29.6	.71	.00
Masking	88			0.30	1, 17.2	.59	.00
BPM	72	0.78****	0.56, 1.00				
BTM	16	0.92**	0.45, 1.40				
Visibility measured	88			0.90	1, 22.4	.35	.00
No	14	0.98***	0.56, 1.39				
Yes	74	0.75****	0.53, 0.97				
Nature of visibility measure	74			1.65	2, 19.4	.22	.00
Objective	57	0.68***	0.44, 0.92				
Subjective	10	0.85**	0.31, 1.40				
Both	7	1.26***	0.66, 1.85				
d'	58	1.04**	0.32, 1.76	8.09	1, 54.1	.006	.20

Note. The regression coefficients for the categorical variables can be interpreted as the mean effect sizes for each category. Model R^2 refers to the proportion of the explained total variance across the three levels using the median sampling variance. k = number of effect sizes in the category; β = regression coefficients; CI = confidence interval; SOA = stimulus onset asynchrony; BPM = backward pattern mask; BTM = backward target mask; d' = sensitivity measure based on signal detection theory.
* $p < .05$. ** $p < .01$. *** $p < .001$. **** $p < .0001$.

(Burnham & Anderson, 1998). The CAICs (k = 88) were 148.9 and 178.6 for the empty and full model, respectively. These values indicate that the full model fit the data no better than the empty model. The backward strategy eliminated the following sequence of nonsignificant moderators: target set size, sample size, prime–target format, masking, number of trials, population, prime format, prime duration, whether or not visibility was measured, and number of target repetitions. We were left with a model containing prime novelty, $F(2, 71.2) = 10.04$, $p = .0001$, category size, $F(1, 21.6) = 26.89$, $p < .0001$, and SOA, β = 0.002, $F(1, 35.3) = 7.00$, $p = .01$. This model (k = 88) had an R^2 of .49. The CAICs (k = 88) were 148.9 and 122.8 for the empty and full model, respectively. These values indicate that the full model performed better than the empty one. Of note, the same final model was obtained irrespective of the starting point for the backward strategy.

In a final step, we added d' to this final model. Three of the four moderators in this model were significant: $F(1, 35.8) = 19.58$, $p < .0001$, for prime novelty, $F(1, 10.7) = 8.76$, $p = .01$, for category size, and β = 0.65, $F(1, 42.6) = 4.15$, $p = .05$, for d'. SOA was no longer significant, $F(1, 13.6) = 1.19$, $p = .29$. This model (k = 58) had an R^2 of .52. The CAICs (k = 58) were 97.2 and 80.5 for the empty and full model, respectively. These values indicate that the full model, which included d', performed better than the empty model. We note that we obtained better fit in a model containing only prime novelty, category size, and d' (excluding SOA). All moderators were significant in this model: $F(1, 37.5) = 18.25$, $p = .0001$, for prime novelty, $F(1, 10.8) = 8.69$, $p = .01$, for category size, and β = 0.65, $F(1, 43.8) = 4.21$, $p = .05$ for d'; the R^2 was .53, and the CAIC 69.9.

Publication Bias and Dependency

We constructed a funnel graph (see Figure 3) to examine publication bias, plotting sample size against effect size. Visual in-

Table 2
Pearson Correlations and Variance Inflation Factors (VIF) Among the Moderators for the Semantic Categorization Conditions

	Correlation																	
Moderator	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	VIF
1. *N*	—	−.25*	.11	.03	.12	−.10	.35**	−.28*	−.55***	−.06	.29*	−.05	.57***	−.42**	.61***	−.17	.07	2.35
2. Population		—	.03	−.07	.03	−.08	−.19	.46***	.33**	−.11	−.09	.003	−.18	.22	−.16	.20	−.01	1.65
3. Prime format			—	.20	.87***	.49***	.26*	−.37**	−.27*	−.03	.34**	.03	−.05	−.05	−.36**	−.58***	−.40**	6.49
4. Prime novelty				—	.27*	.38**	.39**	−.13	−.01	.09	−.12	.25*	−.15	−.04	−.38**	−.07	−.52***	1.36
5. Target format					—	.63***	.46***	−.45***	−.33**	−.03	.12	.22	−.05	.10	−.39**	−.66***	−.47***	7.80
6. Category size						—	.45***	−.34**	−.02	−.10	−.34**	.24	−.36**	.41**	−.63***	−.38*	−.59***	3.28
7. Set size							—	−.65***	−.26*	.19	−.15	.41**	.21	−.17	−.25	−.28	−.18	3.84
8. Target repetitions								—	.49***	−.06	−.10	−.09	−.24	.08	−.07	.42**	.06	3.37
9. Trials									—	−.11	−.14	−.007	−.45***	.28	−.18	.24	−.11	2.59
10. Prime–target format										—	−.09	.41**	−.11	−.19	−.07	−.15	.10	1.48
11. Prime duration											—	−.42**	.28*	−.61***	.32*	−.36*	.09	2.61
12. SOA												—	−.36**	.08	−.22	−.16	.01	1.88
13. Masking													—	−.42***	.88***	/	.17	/
14. Visibility measured														—	/	/	−.35**	/
15. Obj/subj visibility															—	/	.27	/
16. d'																—	.44**	/
17. Effect size																	—	—

Note. To allow the calculation of the VIF, we reduced all categorical moderators to dichotomous variables by eliminating the level with the least observations. The correlations that could not be calculated (marked with /) were the following: the correlation between d' (sensitivity measure based on signal detection theory) and masking, because all studies reporting a d' value used the same masking procedure (backward pattern mask); the correlation between d' and whether or not visibility was measured, because all studies that reported a d' measure used a visibility measure; the correlation between d' and the nature of the visibility measure, because all studies reporting a d' value used an objective visibility test; the correlation between whether or not visibility was measured and the nature of the visibility measure, because all studies reporting an objective or subjective visibility test included a visibility measure. Because of these missing correlations, the VIF for masking, whether or not visibility was measured, and the nature of the visibility measure could not be computed. The VIF for d' was not computed because this would have reduced the number of observations considerably. SOA = stimulus onset asynchrony; obj/subj visibility = nature of the visibility measure (objective/subjective).
* $p < .05$. ** $p < .01$. *** $p < .001$.

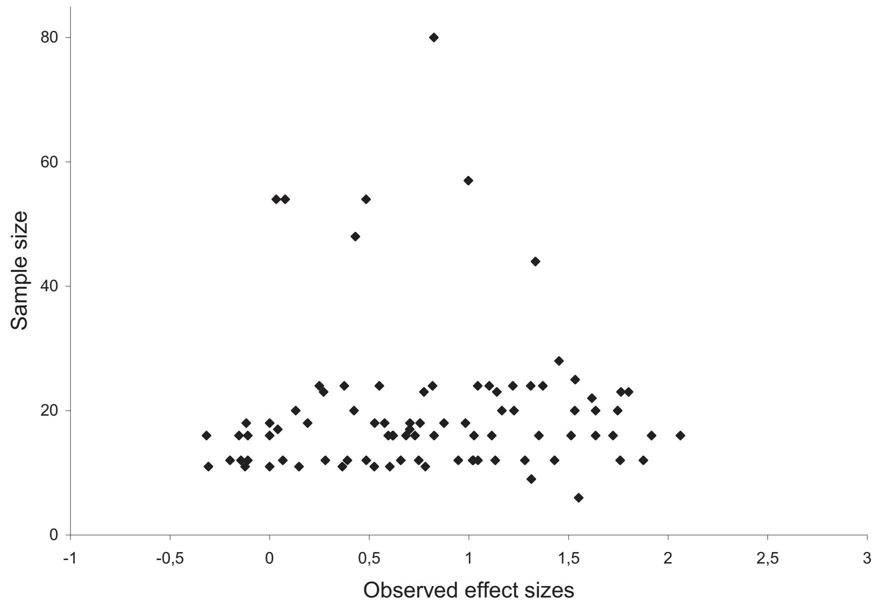

Figure 3. Funnel plot of the semantic categorization conditions.

spection suggests that the graph is symmetrical around the mean effect size (0.80) and does not appear to be heavily skewed. This was confirmed by a nonsignificant correlation between sample size and effect size ($r = .001$, $p = .99$), suggesting that no strong publication bias was present. Additional support for this conclusion is that (a) sample size did not moderate effect sizes, $p = .87$ (i.e., smaller samples did not yield larger effect sizes) and (b) the status of the condition as published or unpublished did not moderate effect sizes; the average effect size for conditions in published studies did not significantly differ from the average effect size for conditions in unpublished studies, $t(14.6) = 1.49$, $p = .16$.

Sensitivity analyses were used to study the impact of dependency among conditions. As explained earlier, some samples contributed more than one effect size to the meta-analysis; 40.9% of the conditions (36 of 88 conditions, indicated in Appendix B) shared a sample with another condition. We constructed two new data sets using random selection procedures such that each condition had an independent sample (no sample contributed more than one effect size). All analyses were conducted again on the two new data sets ($k = 52$). The mean effect sizes for the random effects models were now 0.79 for the first random selection and 0.81 for the second random selection. The regression models with one moderator yielded a pattern of results that was similar to our previous findings, with the same moderators reaching significance (prime format, prime novelty, category size, target format, and d', although the latter was only marginally significant in the second random selection: $p = .08$). As before, none of the two-way interactions were significant and the optimal multiple regression model contained prime novelty, category size, and SOA. These results suggest that the influence of dependency among the effect sizes was limited. When the data were analyzed on data sets in which dependency was eliminated, the pattern of results did not change.

Discussion

We identified 23 studies comprising 88 conditions in which a semantic categorization task was used. The overall mean effect size for the random effects model was 0.80. Significant variance was found at the between-study and the between-condition/within-study levels. In the regression models that contained one moderator, several variables were identified that significantly moderated the effect sizes. We discovered, however, that these regression models were confounded because multicollinearity was present.

We eliminated the multicollinearity and conducted analyses with multiple moderator variables. These analyses showed that the combination of prime novelty, category size, and SOA explained almost half of the variance in effect sizes. Adding the d' measure (thereby reducing the number of studies) improved the model slightly.

First, prime novelty was a strong moderator of the effect sizes: Conditions in which primes and targets were completely different stimuli (i.e., novel primes) showed smaller priming effects than conditions in which primes were also presented as targets (i.e., repeated primes). This observation supports Damian's claim (2001) that priming is enhanced for repeated primes for which nonsemantic S–R links can be formed. We should note, however, that even though priming is diminished when novel primes are used, priming was still significant.

Second, category size moderated the effect sizes: priming was smaller when large rather than small categories were used. This is consistent with research by Forster and colleagues (Forster, 2004; Forster et al., 2003), with the exception that our meta-analysis showed that even though priming was diminished for stimuli from large categories, it was still significant.

Third, SOA moderated the effect sizes: priming effects increased as SOA increased. This observation is consistent with findings by Vorberg et al. (2004) who reported the same relation between response priming and SOA. Nonetheless, the intercept was significant, $\beta = 0.75$, $t(20.6) = 4.45$, $p = .0002$, indicating that significant priming occurs even at very short SOAs.

Finally, the d' measure of prime visibility moderated the effect sizes in those studies that reported d' measures ($k = 58$): Effect sizes decreased as d' measures decreased. This makes intuitive sense: As primes become more visible, they exert a stronger influence, increasing the amount of priming. Our results also show, however, that priming was significant even when prime visibility was 0, indicating that priming can occur in the absence of prime awareness.

Our finding that significant priming occurred under circumstances in which priming was diminished (i.e., novel primes, large categories, and zero prime visibility) is consistent with the claim that primes are processed semantically. Kunde et al. (2003), however, have proposed an alternative nonsemantic hypothesis that might also explain why priming can occur with novel subliminal primes. According to this explanation, participants consciously prepare themselves for a semantic categorization task (e.g., categorize numbers as smaller or larger than 5) by forming *action triggers* for the stimuli that they might receive during the experiment (e.g., numbers between 1 and 9). According to Kunde et al., the fact that only some of these expected stimuli appear as targets (e.g., 1, 4, 6, 9) is irrelevant. Participants will prepare a response to all expected stimuli. If the primes are among the expected set (e.g., 2, 3, 7, 8), participants will have prepared action triggers for them which will facilitate response without the need to process the primes semantically. The formation of an action trigger set should be limited, however, by the number of stimuli and the size of the stimulus category. It seems unlikely that participants could form action triggers for all members of a large category, such as animals or objects. Thus, this theory would not be useful for explaining why we observed significant priming for large categories and large target sets. In order to completely eliminate this alternative, nonsemantic explanation, however, we examined patterns of priming in naming and lexical decision tasks. When participants name targets or make word/nonword judgments, the same action triggers should be formed for related and unrelated primes, since they always evoke the same response.

LEXICAL DECISION AND NAMING

Results

Descriptive Statistics

We identified 32 studies that met all inclusion criteria and used a lexical decision or a naming task in at least 1 condition. In total, 68 conditions were extracted from these 32 studies. All conditions included an academic or nonacademic sample with a mean sample size of 33. A lexical decision task was used in the majority of conditions. Primes and targets were almost always words, and in most conditions, primes and targets were presented in the same format. Primes were presented for an average of 47 ms, with an average SOA of 150 ms. Novel primes were used in almost all conditions. An average of 95 targets was presented per condition, each target usually appearing only once. The average number of trials in a condition was 211. Stimuli from large categories and pattern masks were used in almost all conditions. Only 4 conditions included a d' measure. Appendix A provides descriptive statistics for the 68 lexical decision and naming conditions as a function of potential moderators.

Overall Mean Effect Size and Effect Size Heterogeneity

Appendix C contains the observed effect sizes and corresponding sampling errors for the 68 conditions. Across all conditions, the mean effect size for the random effects model was 0.47 ($k = 68$, 95% CI = 0.36–0.59). Figure 4 graphically displays all observed effect sizes ordered by size and the overall mean effect size with 95% CIs.

A likelihood ratio test comparing models with and without between-study variance showed that significant variance was present at the between-study level, Var = 0.07, $\chi^2(1, k = 68) = 8.6$, $p = .002$, suggesting that moderators of the effect sizes are likely. No significant differences were found between conditions within studies, Var = 0.01, $\chi^2(1, k = 68) = 1.7$, $p = .10$.

We assessed how much of the variance was situated at each of the three levels in the meta-analysis by using the median sampling variance to calculate the total amount of variance. The median amount of sampling variance was 0.06. Thus, the total sum of the between-study variance, the within-study/between-condition variance, and the median sampling variance was 0.14 (0.07 + 0.01 + 0.06). From this, we calculated the percentage of variance situated at the between-study level (49%) and at the within-study/between-condition level (10%). On the basis of the Hunter and Schmidt 75% rule, we can again conclude that the effect sizes are heterogeneous, suggesting that moderators may account for the variability in the effect sizes.

Regression Models With One Moderator

We first examined random effects models in which each moderator was entered separately. We note that a moderator was included only when sufficient observations were available. Four moderators were excluded because of severe restrictions in range: prime format (symbol primes used in only five conditions), prime novelty (repeated primes used in only two conditions), category size (small category used in only two conditions), and d' (only four conditions reported a d' measure). Table 3 provides an overview of the regression models for the remaining 13 moderators.

Sample size. Sample size was a significant moderator of the observed effect sizes, with larger sample sizes associated with smaller effects, $\beta = -0.004$, $F(1, 28.5) = 7.23$, $p = .01$. The proportion of variance explained by sample size was negligible, however.

Target set size. Target set size was a strong moderator of effect sizes, with larger target sets associated with stronger effects, $\beta = 0.002$, $F(1, 18.2) = 17.52$, $p = .0005$. Target set size explained 34% of the variance. The intercept was significant, $\beta = 0.29$, $t(13.7) = 4.73$, $p = .0003$, indicating that significant priming can be expected even when target set sizes are small.

Number of trials. The number of trials in a condition was a moderator of the effect sizes, with larger numbers of trials associated with stronger effects, $\beta = 0.0008$, $F(1, 19) = 10.30$, $p = .005$. The number of trials explained 24% of the variance. The intercept was significant, $\beta = 0.32$, $t(14.8) = 4.45$, $p = .0005$, indicating that priming can be expected even when few trials are used.

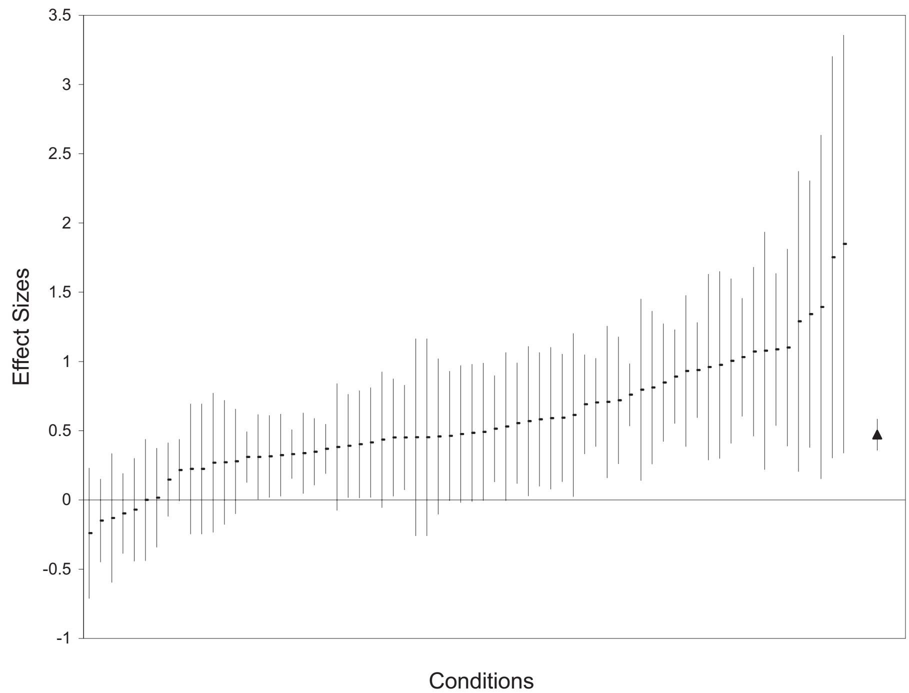

Figure 4. Observed effect sizes for the lexical decision and naming conditions ordered by size of the effect size and the overall mean effect size (indicated by a triangle) with their 95% confidence intervals.

Visibility measured. Whether or not visibility was measured also was a significant moderator of the effect sizes. The average effect size for conditions that included a visibility test was significantly larger than the average effect size for conditions that did not report a visibility test, 0.31 and 0.62, respectively, $t(23.1) = 2.78$, $p = .01$. Significance tests of the regression coefficients showed that the average effect sizes for conditions that used a visibility test and for conditions that did not were both significantly different from 0 (see Table 3). The moderator explained 11% of the variance.

Other moderators had no significant influence on effect sizes. These included population, task, target format, target repetitions, prime–target format, prime duration, SOA, masking, and the nature of the visibility measure. More details are provided in Table 3.

Regression Models With Two-Way Interactions

We examined interactions among the significant moderators (sample size, target set size, number of trials, and whether or not visibility was measured). None of the two-way interactions were significant (p values ranging from .22 to .98). No higher order interactions were investigated for the reasons described earlier.

Multiple Regression Models

Pearson correlations among the moderators (excluding prime format, prime novelty, category size, and d' for reasons previously mentioned) were computed. We transformed categorical variables into dichotomous ones by eliminating the level(s) with the least observations. We eliminated the category "both" in the variable "nature of visibility measure." As a result, 66 of the 68 conditions were retained for analysis. The correlations and VIF values are shown in Table 4. The VIF values exceeded the cutoff for target set size (VIF = 31.50) and number of trials (VIF = 37.72). The substantial overlap between target set size and number of trials makes sense since a larger number of targets is usually presented less often to limit the total number of trials presented to the participants. The number of trials is theoretically the least interesting of the two variables; thus, we decided to omit it from the multiple regression models. After we eliminated number of trials, all remaining VIF values fell well below the cutoff value (VIF range 1.33–3.43). The observed overlap among some variables suggests confounding; thus, care should be taken in interpreting the results from the regression models in which only one moderator was included.

We included only main effects in our multiple regression models because no two-way interactions were significant. We used a backward elimination strategy starting with a model that contained all moderators and then eliminating nonsignificant ones step by step on the basis of the size of their p value. The moderator "nature of the visibility measure" could only be computed on a few effect sizes ($k = 38$); thus, it was not included in the full model. The R^2 for the full model ($k = 57$) was .43; significant moderators were sample size, $\beta = -0.004$, $F(1, 40.1) = 8.18$, $p = .007$, prime duration, $\beta = 0.008$, $F(1, 45) = 7.23$, $p = .01$, and whether or not visibility was measured, $F(1, 13.2) = 10.48$, $p = .006$. The CAICs ($k = 57$) were 46.6 and 76.5 for the empty and full model, respectively. Thus, the full model fit the data no better than the empty model.

The backward strategy eliminated the following sequence of nonsignificant moderators: target format, task, SOA, target repe-

Table 3
Regression Models With One Moderator for Lexical Decision and Naming Conditions

Moderator	k	β	95% CI	F	df	p	Model R^2
N	68	−0.004*	−0.007, −0.001	7.23	1, 28.5	.01	.00
Population	68			1.05	1, 31.0	.31	.01
Academics	57	0.45****	0.32, 0.57				
Nonacademics	11	0.60***	0.33, 0.87				
Task	68			1.88	1, 50.4	.18	.01
Lexical decision	52	0.51****	0.38, 0.63				
Naming	16	0.37***	0.20, 0.55				
Target format	68			0.01	1, 21.8	.91	.00
Symbols	9	0.46**	0.17, 0.75				
Words	59	0.48****	0.35, 0.60				
Set size	60	0.002***	0.001, 0.003	17.52	1, 18.2	.0005	.34
Target repetitions	60	−0.05	−0.12, 0.03	1.53	1, 24.4	.23	.03
Trials	60	0.0008**	0.0003, 0.001	10.30	1, 19.0	.005	.24
Prime–target format	68			1.12	1, 18.6	.30	.00
Same	58	0.50****	0.38, 0.63				
Different	10	0.34*	0.08, 0.61				
Prime duration	68	0.003	−0.002, 0.008	1.11	1, 53.1	.30	.00
SOA	57	0.000008	−0.0006, 0.0006	0.00	1, 23.0	.98	.00
Masking	68			0.02	1, 24.8	.90	.00
BPM	39	0.47****	0.32, 0.62				
BTM	29	0.48****	0.30, 0.66				
Visibility measured	68			7.75	1, 23.1	.01	.11
No	30	0.31***	0.16, 0.46				
Yes	38	0.61****	0.46, 0.76				
Nature of visibility measure	38			0.54	2, 21.6	.59	.00
Objective	27	0.67****	0.47, 0.86				
Subjective	9	0.52**	0.25, 0.78				
Both	2	0.62*	0.12, 1.12				

Note. The regression coefficients for the categorical variables can be interpreted as the mean effect sizes for each category. Model R^2 refers to the proportion of the explained total variance across the three levels using the median sampling variance. k = number of effect sizes in the category; β = regression coefficients; CI = confidence interval; SOA = stimulus onset asynchrony; BPM = backward pattern mask; BTM = backward target mask.
* $p < .05$. ** $p < .01$. *** $p < .001$. **** $p < .0001$.

titions, masking, prime–target format, and population. We were left with a model containing sample size, β = −0.003, $F(1, 16.11) = 5.79$, $p = .03$, whether or not visibility was measured, $F(1, 13.4) = 9.46$, $p = .009$, prime duration, β = 0.007, $F(1, 35.2) = 8.24$, $p = .007$, and target set size, β = 0.002, $F(1, 17.7) = 9.18$, $p = .007$. This model ($k = 60$) had an R^2 of .43. The CAICs ($k = 60$) were 51.3 and 58.7 for the empty and full model, respectively. Thus, the final model fit the data no better than the empty model. Of note, the same final model was obtained irrespective of the starting point for the backward strategy.

Publication Bias and Dependency

We constructed a funnel graph (see Figure 5). Visual inspection suggests that the graph is imperfectly symmetrical around the mean effect size (0.47). This is confirmed by the significant correlation between sample size and effect size ($r = -.29$, $p = .02$), suggesting that publication bias is present. The existence of publication bias is also supported by evidence that (a) sample size moderated effect sizes, $p = .01$, (i.e., smaller samples yielded larger effect sizes) and (b) the status of the condition as published or unpublished did not moderate the observed effect sizes, $t(20.5) = 1.45$, $p = .16$; nonetheless, the average effect size was significant in published conditions, 0.50, $t(21.2) = 8.24$, $p < .0001$, but not in unpublished conditions, 0.22, $t(20.5) = 1.21$, $p = .24$.

Sensitivity analyses were again used to study the impact of dependency among conditions. Ten of the 68 conditions (14.7%, indicated in Appendix C) shared a sample with another condition. We conducted the analyses again on two new data sets that were constructed as described earlier ($k = 58$). The overall mean effect sizes for the random effects models were now 0.47 for the first random selection and 0.46 for the second random selection. The regression models with one moderator yielded the same pattern of results as they did before, with the same significant moderators (sample size, target set size, prime duration, and whether or not prime visibility was measured). As before, none of the two-way interactions were significant, and the optimal multiple model contained sample size, whether or not visibility was measured, prime duration, and target set size. These results suggest that the influence of dependency among the effect sizes was limited. When all dependency was eliminated, the pattern of results did not change.

Discussion

We identified 32 studies comprising 68 conditions in which either a lexical decision or naming task was used. The overall mean effect size for the random effects model was 0.47. Significant variance was found at the between-study level. In the regression models that contained one moderator several variables were identified that significantly moderated the effect sizes. It was clear,

Table 4
Pearson Correlations and Variance Inflation Factors (VIF) Among the Moderators for the Lexical Decision and Naming Conditions

	Correlation														
Moderator	1	2	3	4	5	6	7	8	9	10	11	12	13	14	VIF
1. *N*	—	−.21	.04	−.002	−.02	−.11	−.08	.14	.006	.09	−.09	−.11	.49*	−.28*	1.34
2. Population		—	−.24	.17	.66***	−.16	.73***	−.18	−.17	−.12	−.12	.22	−.31	.12	2.94
3. Task			—	−.50***	−.30*	.19	−.36**	.65***	.07	.07	−.002	−.19	.09	−.25*	2.19
4. Target format				—	.22	−.07	.23	−.69***	−.24*	.11	.17*	.17	−.29	.09	2.60
5. Set size					—	−.27*	.97***	−.26*	−.15**	−.14	−.17	.38*	−.26	.27*	31.50
6. Target repetitions						—	−.15	.22	−.18	.18	−.21	−.01	−.07	−.28*	1.88
7. Trials							—	−.29*	−.21	−.12	−.20	.41**	−.31	.24	37.72
8. Prime–target format								—	.12	.17	−.29*	−.04	.14	−.18	3.45
9. Prime duration									—	−.46***	.31*	−.50***	.28	−.25*	2.32
10. SOA										—	−.54***	.32*	−.41*	.30*	2.89
11. Masking											—	−.29*	.45**	−.13	2.74
12. Visibility measured												—	/	.41**	2.01
13. Obj/subj visibility													—	−.22	/
14. Effect size														—	—

Note. To allow the calculation of the VIF, we reduced all categorical moderators to dichotomous variables by eliminating the level with the least observations. The correlation between whether or not visibility was measured and the nature of the visibility measure (marked with /) could not be calculated, because all studies reporting an objective or subjective visibility test included a visibility measure. Because of this missing correlation, the VIF for the nature of the visibility measure could not be computed. SOA = stimulus onset asynchrony; obj/subj visibility = the nature of the visibility measure (objective/subjective).
* $p < .05$. ** $p < .01$. *** $p < .001$.

however, that these models were confounded because multicollinearity was present.

We eliminated the multicollinearity and conducted analyses with multiple moderator variables. These analyses showed that the combination of sample size, target set size, prime duration, and whether or not prime visibility was measured explained almost half of the variance in effect sizes.

First, with respect to sample size, priming effects were diminished as sample size increased. This indicates that publication bias is likely because larger studies reported significantly smaller effect sizes than smaller studies.

Second, target set size moderated the effect sizes: Priming increased as set size increased. This concurs with findings reported by Abrams (2008) and Kiesel et al. (2006). They found priming effects

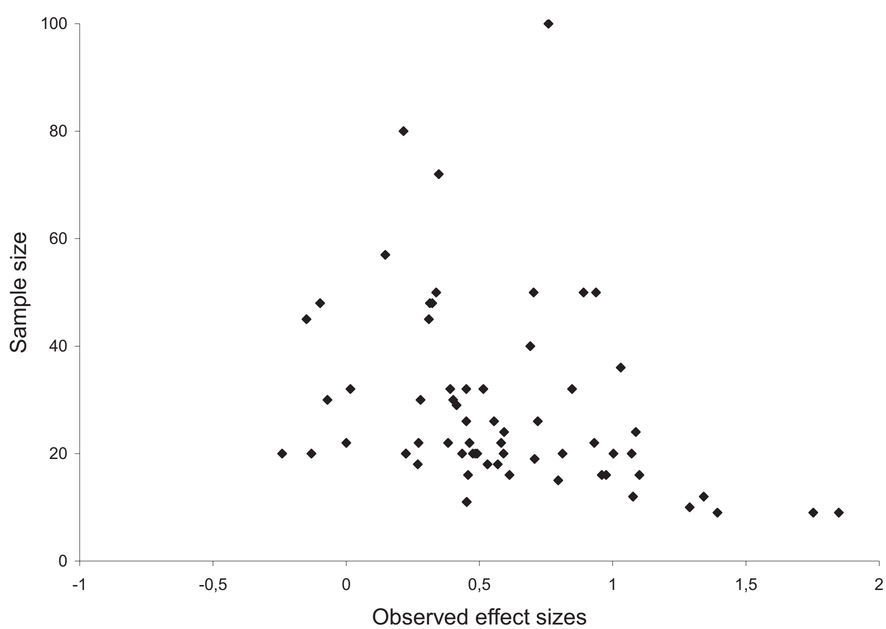

Figure 5. Funnel plot of the lexical decision and naming conditions.

using novel primes from a large category but only when they used large target sets. Decreased priming in conditions with small target sets may be due to the development of S–R mappings. When the set size is small, each target is likely to have at least one feature that distinguishes it from the other targets. Participants may discover these unique features and use them to complete the task. This would minimize the need to process the targets semantically (see also Van den Bussche & Reynvoet, in press). An alternative means of developing S–R mappings would be to store each target along with its response in short-term memory. This could be done with small target sets but not with large ones because only a few items can be stored in short-term memory (Roelofs, 2001). Nonetheless, our results showed significant effect sizes even when target set size approached zero.

Third, prime duration also moderated effect sizes: They decreased as prime duration decreased. This corresponds to the findings of several studies (e.g., Holcomb et al., 2005; Klauer et al., 2007). Nonetheless, we found significant effect sizes even when prime duration was very short.

Finally, whether or not prime visibility was measured moderated the effect sizes: Priming was larger in those conditions that assessed prime visibility relative to those that reported no prime visibility measure. Priming was significant, however, even when a visibility measure was not reported.

We note that four variables—prime format, prime novelty, category size, and d'—could not be included in the meta-analysis because some levels of these variables were underrepresented in the literature. The exclusion of several important moderators may have diminished the explanatory power of the models. Furthermore, it seems likely that publication bias resulted in overestimation of the overall mean effect size. Publication bias may have been caused by several factors. First, the number of effect sizes from unpublished studies was relatively small (12%). Second, almost all studies published in the 1980s were included in the lexical decision and naming meta-analysis. These older studies yielded larger effect sizes than more recent studies but were also more likely to have some methodological flaws (Holender, 1986). Third, the reluctance to publish nonsignificant priming results may have been stronger in earlier than in later years. Indeed, all signs of publication bias disappeared when we conducted the analyses again using only the unpublished conditions and the conditions published after 1998 (publication year of Dehaene et al.). Therefore, caution is warranted in interpreting the lexical decision and naming results because the influence of several potentially important variables could not be tested and because publication bias likely influenced the results.

General Discussion

The aim of our meta-analyses was twofold. First, we wanted to assess the magnitude of subliminal priming in the literature. We conducted separate analyses of semantic categorization and lexical decision/naming conditions to examine whether subliminal information is processed semantically. Response congruency does not bias lexical decision and naming tasks as it does semantic categorization tasks. Second, we wanted to assess the influence of several potential moderators of subliminal priming in an attempt to distinguish the underlying mechanisms of response priming and true semantic priming (Kiefer, 2007).

Our meta-analyses showed significant priming in studies conducted between 1983 and 2006. Priming was significant for both semantic categorization and lexical decision and naming conditions (as indicated by strong positive overall effect sizes). Priming effects in semantic categorization conditions were moderated primarily by prime novelty, category size, SOA, and d'. Priming effects in lexical decision and naming conditions were moderated primarily by sample size, target set size, prime duration, and whether or not prime visibility was assessed. We found strong support for the claim that subliminal primes can be processed semantically. Significant priming was observed in the context of lexical decision and naming, tasks that are unconfounded by response effects. Moreover, we observed priming in semantic categorization even in those circumstances in which priming is difficult to explain by means other than semantic analysis of the primes (novel primes, large stimulus categories).

Our results also showed that priming was larger when the formation of S–R mappings was possible. Priming in semantic categorization conditions, where it is biased by S–R effects, is larger than priming in lexical decision and naming conditions. Furthermore, in semantic categorization conditions, stronger priming was reported when primes were repeated and small categories were used, circumstances in which S–R mappings can easily be formed. These findings are consistent with Damian's (2001) and Abrams and Greenwald's (2000) claims that S–R mappings can lead to enhanced priming. S–R mappings alone cannot explain all of our results, however. Priming was significant even when the influence of S–R mapping was minimized or eliminated (e.g., in lexical decision and naming tasks and semantic categorization with novel primes and/or large categories).

These findings reconcile both reigning theories of subliminal priming: Both can explain the data, depending on the task context. Automatic S–R mappings clearly boost priming effects when the task context enables the S–R associations to be formed (Abrams & Greenwald, 2000; Damian, 2001). When the opportunity to form S–R mappings is minimized, however, subliminal priming can occur by means of semantic processing of the primes (e.g., Klauer et al., 2007; Van den Bussche, Notebaert, & Reynvoet, in press; Van den Bussche & Reynvoet, 2007). This latter claim is consistent with evidence from recent electrophysiological and brain imaging studies, which show that subliminal primes activate cerebral regions that are associated with semantic processing. For example, Kiefer and Brendel (2006) reported that the N400 (an event-related potential component that is sensitive to semantic integration) was modulated by subliminal semantic priming in a lexical decision task, even when prime awareness was controlled carefully. Devlin, Jamison, Matthews, and Gonnerman (2004) compared masked prime–target pairs that were related (e.g., *idea–notion*) to unrelated pairs and found reduced neural activity to related pairs in the left middle temporal gyrus, a region thought to be involved in semantic processing of words and objects.

Our findings revealed some of the factors that determine when S–R mappings are likely to develop. These factors are important in designing subliminal priming studies. If one aims to investigate the semantic processing of subliminal information, a design should be chosen that excludes the influence of S–R effects as much as possible because they can confound the results.

Our findings also revealed that the overall effect size seems to be smaller in lexical decision and naming than in semantic categorization. Moreover, different variables moderate priming in semantic categorization (prime novelty, category size, SOA, and the d' measure) than in lexical decision and naming (sample size, target set size, prime duration, and whether or not prime visibility is measured).

These findings affirm our decision to analyze these two kinds of tasks separately. One explanation for observed differences between the tasks is that differential mechanisms underlie priming in semantic categorization and in lexical decision and naming. In semantic categorization, priming is enhanced for repeated primes and small categories. S–R mappings provide a reasonable explanation for priming in these conditions. In lexical decision and naming, priming is enhanced for large target sets. A large target set may force participants to semantically process the targets (because they can no longer form S–R mappings), allowing the primes to exert a larger influence on the targets and consequently evoke larger priming. We stress, however, that we were unable to enter the exact same moderators in the two meta-analyses because some had insufficient variability in the lexical decision and naming conditions. Furthermore, the lexical decision and naming data appear to be influenced by publication bias. Therefore, caution is warranted in directly comparing results from our two meta-analyses.

As mentioned earlier, the lexical decision and naming studies often included the same combinations of moderators: novel primes, word primes, large category stimuli, and large target sets. This creates gaps in our understanding of the factors that moderate priming in these tasks. Future research should address these missing combinations of moderators. For example, our understanding of how S–R mappings influence subliminal priming would be enhanced by lexical decision and naming studies using repeated primes and/or small categories. Do these factors enhance priming in lexical decision and naming (unbiased by response effects) as they do semantic categorization?

Finally, our meta-analyses clearly indicated the importance of including a visibility test in subliminal priming experiments. Even more preferable are studies that include a sensitivity measure such as d' to assess prime awareness, as indicated by the significant effect of d' on priming in the semantic categorization studies. Conclusions about the impact of unconscious information on behavior can be made only when we are fairly confident that the information was processed unconsciously. Thus, some assessment of prime awareness is indispensable. Only four lexical decision and naming studies included a d' measure. We hope that our meta-analyses encourage researchers to include thorough visibility measures in future studies.

Our meta-analyses have a few important limitations. First, we were unable to compute higher order interactions among moderators because of gaps in literature and confounding among several moderators (e.g., almost two thirds of the studies with a large category also used novel primes). Furthermore, we did not obtain significant two-way interactions among the moderators. Of course, it may be that these interactions simply do not exist. However, it also seems plausible that the absence of interaction effects was due to severe restrictions in range for some moderators since recent studies have provided some evidence suggesting that interactions among certain moderators are present (e.g., interaction between category size and target set size; Kiesel et al., 2006). Second, dependency among moderators did not affect priming in either meta-analysis; however, publication bias affected our analysis of the lexical decision and naming data. Publication bias was not a significant moderator of priming in semantic categorization. Nonetheless, we cannot completely rule out the possibility that these issues influenced the results slightly. Third, prime awareness may not have been sufficiently controlled in some of the conditions in our analyses, such as in conditions in which relatively long prime durations or SOAs were used, conditions that did not report a visibility test, and conditions with an inefficient masking procedure. However, some of these variables had no significant impact on priming in any of our analyses (masking and the nature of the visibility test). For the other variables, we found significant priming under the most restrictive conditions. These variables included prime duration, whether or not prime visibility was assessed, and d'.

Can subliminally presented information influence our behavior? Our meta-analyses of research conducted between 1983 and 2006 indicate that the answer to this question is yes. Furthermore, our quantitative review confirms that subliminal information can be processed semantically. However, our study also shows that nonsemantic processing of subliminal information can boost priming effects when the experimental context allows for the opportunity to form nonsemantic S–R mappings.

References

References marked with an asterisk indicate studies included in the meta-analysis.

Abrams, R. L. (2008). Influence of category size and target set size on unconscious priming by novel words. *Experimental Psychology, 55,* 189–194.

Abrams, R. L., & Greenwald, A. G. (2000). Parts outweigh the whole (word) in unconscious analysis of meaning. *Psychological Science, 11,* 118–124.

*Alameda, J. R., Cuetos, F., & Brysbaert, M. (2003). The number 747 is named faster after seeing Boeing than after seeing Levi's: Associative priming in the processing of multidigit Arabic numerals. *The Quarterly Journal of Experimental Psychology A, 56,* 1009–1019.

*Bajo, M. T., Puerta-Melguizo, M. C., & Macizo, P. (2003). The locus of semantic interference in picture naming. *Psicológica, 24,* 31–55.

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

Begg, C. B. (1994). Publication bias. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 399–410). New York: Russell Sage Foundation.

*Bodner, G. A., & Dypvik, A. T. (2005). Masked priming of number judgments depends on prime validity and task. *Memory & Cognition, 33,* 29–47.

*Bodner, G. E., & Masson, M. E. J. (2003). Beyond spreading activation: An influence of relatedness proportion on masked semantic priming. *Psychonomic Bulletin & Review, 10,* 645–652.

*Bourassa, D. C., & Besner, D. (1998). When do nonwords activate semantics? Implications for models of visual word recognition. *Memory & Cognition, 26,* 61–74.

*Brown, M., & Besner, D. (2002). Semantic priming: On the role of awareness in visual word recognition in the absence of an expectancy. *Consciousness and Cognition, 11,* 402–422.

*Bueno, S., & Frenck-Mestre, C. (2002). Rapid activation of the lexicon: A further investigation with behavioral and computational results. *Brain and Language, 81,* 120–130.

Burnham, K. P., & Anderson, D. R. (1998). *Model selection and inference: A practical information-theoretic approach.* New York: Springer–Verlag.

*Carr, T. H., & Dagenbach, D. (1990). Semantic priming and repetition priming from masked words: Evidence for a center-surround attentional mechanism in perceptual recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 341–350.

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

*Damian, M. F. (2001). Congruity effects evoked by subliminally presented primes: Automaticity rather than semantic processing. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 154–165.

*Dehaene, S., Naccache, L., Le Clec'H, G., Koechlin, E., Mueller, M., Dehaene-Lambertz, G., et al. (1998, October 29). Imaging unconscious semantic priming. *Nature, 395,* 597–600.

*Dell'Acqua, R., & Grainger, J. (1999). Unconscious semantic priming from pictures. *Cognition, 73,* 1–15.

*Devlin, J. T., Jamison, H. L., Gonnerman, L. M., & Matthews, P. M. (2006). The role of the posterior fusiform gyrus in reading. *Journal of Cognitive Neuroscience, 18,* 911–922.

*Devlin, J. T., Jamison, H. L., Matthews, P. M., & Gonnerman, L. M. (2004). Morphology and the internal structure of words. *Proceedings of the National Academy of Sciences of the United States of America, 101,* 14984–14988.

Draine, S., & Greenwald, A. (1998). Replicable unconscious semantic priming. *Journal of Experimental Psychology: General, 127,* 286–303.

*Finkbeiner, M., & Caramazza, A. (2006). Now you see it, now you don't: On turning semantic interference into facilitation in a Stroop-like task. *Cortex, 42,* 790–796.

*Forster, K. I. (2004). Category size effects revisited: Frequency and masked priming effects in semantic categorization. *Brain and Language, 90,* 276–286.

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 680–698.

*Forster, K. I., & Hector, J. (2005). [Semantic priming as a function of SOA]. Unpublished raw data.

*Forster, K. I., Mohan, K., & Hector, J. (2003). The mechanics of masked priming. In: S. Kinoshita & S. J. Lupker (Eds.), *Masked priming: The state of the art* (pp. 3–37). New York: Psychology Press.

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

*Frost, R., Forster, K. I., & Deutsch, A. (1997). What can we learn from the morphology of Hebrew? A masked-priming investigation of morphological representation. *Journal of Experimental Psychology: Learning, Memory, and Cognitio*n, *23,* 829–856.

Gibbons, R. D., Hedeker, D. R., & Davis, J. M. (1993). Estimation of effect size from a series of experiments involving paired comparisons. *Journal of Educational Statistics, 18,* 271–279.

Glaser, W. R. (1992). Picture naming. *Cognition, 42,* 61–105.

Grainger, J., & Frenck-Mestre, C. (1998). Masked priming by translation equivalents in bilinguals. *Language and Cognitive Processes, 13,* 601–623.

Greenhouse, J. B., & Iyengar, S. (1994). Sensitivity analysis and diagnostics. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 383–398). New York: Russell Sage Foundation.

Greenwald, A. G., Abrams, R. L., Naccache, L., & Dehaene, S. (2003). Long-term semantic memory versus contextual memory in unconscious number processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 235–247.

Greenwald, A. G., Draine, S. C., & Abrams, R. L. (1996, September 20). Three cognitive markers of unconscious semantic activation. *Science, 273,* 1699–1702.

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

*Grossi, G. (2006). Relatedness proportion effects on masked associative priming: An ERP study. *Psychophysiology, 43,* 21–30.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6,* 107–128.

Hedges, L. V. (1982). Estimation of effect size from a series of independent experiments. *Psychological Bulletin, 92,* 490–499.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* Orlando, FL: Academic Press.

*Hines, D. (1993). Effect of masked picture primes on semantic priming for easy- and difficult-to-name words. *The Journal of General Psychology, 120,* 149–165.

*Hines, D., Czerwinski, M., Sawyer, P. K., & Dwyer, M. (1986). Automatic semantic priming: Effect of category exemplar level and word association level. *Journal of Experimental Psychology: Human Perception and Performance, 12,* 370–379.

Holcomb, P. J., Reder, L., Misra, M., & Grainger, J. (2005). The effects of prime visibility on ERP measures of masked priming. *Cognitive Brain Research, 24,* 155–172.

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.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings.* Newbury Park, CA: Sage.

*Kamphuis, J. H., De Ruiter, C., Janssen, B., & Spiering, M. (2005). Preliminary evidence for an automatic link between sex and power among men who molest children. *Journal of Interpersonal Violence, 20,* 1351–1365.

*Kemp-Wheeler, S. M., & Hill, A. B. (1988). Semantic priming without awareness: Some methodological considerations and replications. *The Quarterly Journal of Experimental Psychology, A, 40,* 671–692.

*Kiefer, M. (2002). The N400 is modulated by unconsciously perceived masked words: Further evidence for an automatic spreading activation account of N400 priming effects. *Cognitive Brain Research, 13,* 27–39.

Kiefer, M. (2007). Top-down modulation of unconscious "automatic" processes: A gating framework. *Advances in Cognitive Psychology, 3,* 289–306.

*Kiefer, M., & Brendel, D. (2006). Attentional modulation of unconscious "automatic" processes: Evidence from event-related potentials in a masked priming paradigm. *Journal of Cognitive Neuroscience, 18,* 184–198.

*Kiefer, M., & Spitzer, M. (2000). Time course of conscious and unconscious semantic brain activation. *Cognitive Neuroscience, 11,* 2401–2407.

*Kiesel, A., Kunde, W., Pohl, C., & Hoffmann, J. (2006). Priming from novel masked stimuli depends on target set size. *Advances in Cognitive Psychology, 2,* 37–45.

*Kinoshita, S., Mozer, M. C., & Forster, K. I. (2006). *Modulation of congruence effects with masked primes in the parity decision task.* Manuscript submitted for publication.

Klauer, C. K., Eder, A. B., Greenwald, A. G., & Abrams, R. L. (2007). Priming of semantic classifications by novel subliminal prime words. *Consciousness and Cognition, 16,* 63–83.

Klinger, M. R., Burton, P. C., & Pitts, G. S. (2000). Mechanisms of unconscious priming: I. Response competition, not spreading activation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 441–455.

*Koechlin, E., Naccache, L., Block, E., & Dehaene, S. (1999). Primed numbers: Exploring the modularity of numerical representations with masked and unmasked semantic priming. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 1882–1905.

Kouider, S., & Dehaene, S. (2007). Levels of processing during non-conscious perception: A critical review of visual masking. *Philosophical Transactions of the Royal Society B, 362,* 857–875.

*Kunde, W., Kiesel, A., & Hoffmann, J. (2003). Conscious control over the content of unconscious cognition. *Cognition, 88,* 223–242.

*Kunde, W., Kiesel, A., & Hoffmann, J. (2005). On the masking and disclosure of unconscious elaborate processing: A reply to Van Opstal, Reynvoet, and Verguts (2005). *Cognition, 97,* 99–105.

LeDoux, J. E. (1996). *The emotional brain: The mysterious underpinnings of emotional life.* New York: Simon & Schuster.

Light, R. J., & Pillemer, D. B. (1984). *Summing up: The science of reviewing research.* Cambridge, MA: Harvard University Press.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis.* Newbury Park, CA: Sage.

Littell, R. C., Milliken, G. A., Stroup, W. W., Wolfinger, R. D., & Schabenberger, O. (2006). *SAS system for mixed models* (2nd ed.). Cary, NC: SAS Institute.

Lucas, M. (2000). Semantic priming without association: A meta-analytic review. *Psychonomic Bulletin & Review, 7,* 618–630.

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

*McBride, K. A., Kripps, J. H., & Forster, K. I. (2003). [Phonological and semantic priming in deaf and hearing subjects]. Unpublished raw data.

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

Merikle, P. M., & Reingold, E. M. (1992). Measuring unconscious perceptual processes. In R. F. Bornstein & T. S. Pittman (Eds.), *Perception without awareness* (pp. 55–80). New York: Guilford Press.

Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods, 7,* 105–125.

*Naccache, L., & Dehaene, S. (2001). Unconscious semantic priming extends to novel unseen stimuli. *Cognition, 80,* 223–237.

O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity, 41,* 673–690.

*O'Seaghdha, P. G., & Marin, J. W. (1997). Mediated semantic-phonological priming: Calling distant relatives. *Journal of Memory and Language, 36,* 226–252.

*Perea, M., & Gotor, A. (1997). Associative and semantic priming effects occur at very short stimulus-onset asynchronies in lexical decision and naming. *Cognition, 62,* 223–240.

*Perea, M., & Lupker, S. J. (2003). Does judge activate COURT? Transposed-letter similarity effects in masked associative priming. *Memory & Cognition, 31,* 829–841.

*Perea, M., & Rosa, E. (2002a). Does the proportion of associatively related pairs modulate the associative priming effect at very brief stimulus-onset asynchronies? *Acta Psychologica, 110,* 103–124.

*Perea, M., & Rosa, E. (2002b). The effects of associative and semantic priming in the lexical decision task. *Psychological Research, 66,* 180–194.

Perfetti, C. A., & Tan, L. H. (1998). The time course of graphic, phonological, and semantic activation in visual Chinese character identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 101–118.

*Pohl, C., Kiesel, A., Kunde, W., & Hoffmann, J. (2004). [Impact of target size and concept category on the effect of masked primes]. Unpublished raw data.

*Pohl, C., Kiesel, A., Kunde, W., & Hoffmann, J. (2005). [Impact of target size and concept category on the effect of masked primes]. Unpublished raw data.

*Pohl, C., Kiesel, A., Kunde, W., & Hoffmann, J. (2006). *Early and late selection in unconscious information processing.* Manuscript submitted for publication.

Raudenbush, S. W., & Bryk, A. S. (1985). Empirical Bayes meta-analysis. *Journal of Educational Statistics, 10,* 78–98.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models* (2nd ed.). London: Sage.

*Reynvoet, B., Caessens, B., & Brysbaert, M. (2002). Automatic stimulus–response associations may be semantically mediated. *Psychonomic Bulletin & Review, 9,* 107–112.

Reynvoet, B., Gevers, W., & Caessens, B. (2005). Unconscious primes activate motor codes through semantics. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 991–1000.

Roelofs, A. (2001). Set size and repetition matter: Comment on Caramazza and Costa (2000). *Cognition, 80,* 291–298.

Rosenthal, R. (1991). *Meta-analytical procedures for social research.* Newbury Park, CA: Sage.

*Rouibah, A., Tiberghien, G., & Lupker, S. J. (1999). Phonological and semantic priming: Evidence for task-independent effects. *Memory & Cognition, 27,* 422–437.

*Ruz, M., Madrid, E., Lupiáñez, J., & Tudela, P. (2003). High density ERP indices of conscious and unconscious semantic priming. *Cognitive Brain Research, 17,* 719–731.

*Sassenberg, K., & Moskowitz, G. B. (2005). Don't stereotype, think different! Overcoming automatic stereotype activation by mindset priming. *Journal of Experimental Social Psychology, 41,* 506–514.

*Sebba, D., & Forster, K. I. (1989). [Mediated priming with masked primes]. Unpublished raw data.

Sidis, B. (1898). *The psychology of suggestion.* New York: Appleton.

Snijders, T. A. B., & Bosker, R. J. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling.* London: Sage.

Tan, L. H., Hoosain, R., & Peng, D. (1995). Role of early presemantic phonological code in Chinese character identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 43–54.

Tan, L. H., Hoosain, R., & Siok, W. K. (1996). Activation of phonological codes before access to character meaning in written Chinese. *Journal of Experimental Psychology: Learning, Memory and Cognition, 22,* 865–882.

Theios, J., & Amrhein, P. C. (1989). Theoretical analysis of the cognitive processing of lexical and pictorial stimuli: Reading, naming, and visual and conceptual comparisons. *Psychological Review, 96,* 5–24.

*Théoret, H., Halligan, E., Kobayashi, M., Merabet, L., & Pascual-Leone, A. (2004). Unconscious modulation of motor cortex excitability revealed with transcranial magnetic stimulation. *Experimental Brain Research, 155,* 261–264.

Van den Bussche, E., Notebaert, K., & Reynvoet, B. (in press). Masked primes can be genuinely semantically processed: A picture prime study. *Experimental Psychology.*

*Van den Bussche, E., & Reynvoet, B. (2006). [The influence of attention on masked priming effects]. Unpublished raw data.

Van den Bussche, E., & Reynvoet, B. (2007). Masked priming effects in semantic categorization are independent of category size. *Experimental Psychology, 54,* 225–235.

Van den Bussche, E., & Reynvoet, B. (in press). The mechanisms of masked priming: A state of the art. In E. B. Hartonek (Ed.), *Experimental psychology research trends.* New York: Nova Science.

Van den Noortgate, W., & Onghena, P. (2003). Multilevel meta-analysis: A comparison with traditional meta-analytical procedures. *Educational and Psychological Measurement, 63,* 765–790.

*Van Opstal, F., Reynvoet, B., & Verguts, T. (2005a). How to trigger elaborate processing? A comment on Kunde, Kiesel, and Hoffmann (2003). *Cognition, 97,* 89–97.

*Van Opstal, F., Reynvoet, B., & Verguts, T. (2005b). Unconscious semantic categorization and mask interactions: An elaborate response to Kunde et al. (2005). *Cognition, 97,* 107–113.

Vorberg, D., Mattler, U., Heinecke, A., Schmidt, T., & Schwarzbach, J. (2004). Invariant time-course of priming with and without awareness. In C. Kaernbach, E. Schröger, & H. Müller (Eds.), *Psychophysics beyond sensation: Laws and invariants of human cognition* (pp. 273–290). Hillsdale, NJ: Erlbaum.

*Williams, J. N. (1996). Is automatic priming semantic? *European Journal of Cognitive Psychology, 8,* 113–161.

*Zhou, X., Marslen-Wilson, W., Taft, M., & Shu, H. (1999). Morphology, orthography, and phonology in reading Chinese compound words. *Language and Cognitive Processes, 14,* 525–565.

Ziegler, J. C., Ferrand, L., Jacobs, A. M., Rey, A., & Grainger, J. (2000). Visual and phonological codes in letter and word recognition: Evidence from incremental priming. *The Quarterly Journal of Experimental Psychology A, 53,* 671–692.

(*Appendixes follow*)

Appendix A

Operationalization and Descriptive Statistics for Potential Moderators for the Semantic Categorization and for Lexical Decision and Naming Conditions Separately

			Descriptive statistics	
Moderator	Value	Coding description and criteria	Semantic categorization conditions	Lexical decision and naming conditions
Population	1 = academics 2 = nonacademics	Categorical variable representing whether the sample consisted of academics (e.g., university students) or nonacademics.	k = 88 Academics k = 63 Nonacademics k = 25	k = 68 Academics k = 57 Nonacademics k = 11
Sample size	Continuous	Continuous variable representing the sample size of the condition.	k = 88 M = 20, SD = 12.0 Range = 6–80	k = 68 M = 33, SD = 26.7 Range = 9–132
Task	1 = lexical decision 2 = naming	Categorical variable representing whether a lexical decision task or a naming task was used.		k = 68 Lexical decision k = 52 Naming k = 16
Prime format	1 = symbols 2 = words 3 = both	Categorical variable representing whether the prime stimuli were symbols (i.e., digits, letters, pictures or Chinese symbols), words (i.e., words or number words), or both (i.e., mixture of digits and number words).	k = 88 Symbols k = 32 Words k = 38 Both k = 18	k = 68 Symbols k = 5 Words k = 63
Prime novelty	1 = repeated primes 2 = novel primes 3 = both	Categorical variable representing whether the prime stimuli were repeated primes (i.e., primes that were also presented as targets), novel primes (i.e., primes that were never presented as targets), or both.	k = 88 Repeated primes k = 40 Novel primes k = 44 Both k = 4	k = 68 Repeated primes k = 2 Novel primes k = 66
Target format	1 = symbols 2 = words 3 = both	Categorical variable representing whether the target stimuli were symbols (i.e., digits, letters, pictures, or Chinese symbols), words (i.e., words or number words), or both (i.e., mixture of digits and number words).	k = 88 Symbols k = 36 Words k = 38 Both k = 14	k = 68 Symbols k = 9 Words k = 59
Category size	1 = small 2 = large	Categorical variable representing whether the stimuli came from a large category (i.e., animals, objects, positive and negative words, adjectives, or a combination of various stimulus categories) or a small category (i.e., months, numbers below 10, farm animals, types of dogs, body parts, letters, or words related to power or sex).	k = 88 Small k = 44 Large k = 44	k = 60 Small k = 2 Large k = 58
Set size	Continuous	Continuous variable representing the number of unique relevant targets presented to the participants (i.e., excluding nonword targets).	k = 88 M = 21, SD = 23.7 Range = 4–90	k = 60 M = 95, SD = 95.0 Range = 6–320
Target repetitions	Continuous	Continuous variable representing the number of times the targets were repeated during the experiment.	k = 88 M = 68, SD = 64.6 Range = 1–192	k = 60 M = 1, SD = 1.2 Range = 1–8
Trials	Continuous	Continuous variable representing the total number of trials the participants received during the experiment.	k = 88 M = 485, SD = 274.7 Range = 40–1,536	k = 60 M = 211, SD = 217.6 Range = 30–800
Prime–target format	1 = same 2 = different	Categorical variable representing whether primes and targets were presented in the same format (i.e., both were symbols, words, or mixed) or in a different format (e.g., primes were symbols and targets were words or vice versa).	k = 88 Same k = 80 Different k = 8	k = 68 Same k = 58 Different k = 10
Prime duration	Continuous	Continuous variable representing how long the primes were presented to the participants (in milliseconds).	k = 88 M = 42, SD = 11.2 Range = 10–72	k = 68 M = 47, SD = 17.9 Range = 10–84
SOA	Continuous	Continuous variable representing how long the stimulus onset asynchrony (SOA) was (i.e., time interval between prime onset and target onset, in milliseconds).	k = 88 M = 106, SD = 65.3 Range = 41–500	k = 57 M = 150, SD = 184.4 Range = 33–784

(*table continues*)

Table (*continued*)

Moderator	Value	Coding description and criteria	Descriptive statistics: Semantic categorization conditions	Descriptive statistics: Lexical decision and naming conditions
Masking	1 = BPM 2 = BTM	Categorical variable representing whether the condition used a backward pattern mask (BPM; i.e., a pattern mask presented after the prime) or a backward target mask (BTM; i.e., target presented immediately after the prime).	k = 88 BPM k = 72 BTM k = 16	k = 68 BPM k = 39 BTM k = 29
Visibility measured	0 = no 1 = yes	Dichotomous variable representing whether or not the visibility of the primes was measured in the condition.	k = 88 No k = 14 Yes k = 74	k = 68 No k = 30 Yes k = 38
Nature of visibility measure	1 = objective 2 = subjective 3 = both	Categorical variable representing how the visibility of the primes was measured in the condition: either objectively (i.e., an objective test of prime visibility), subjectively (i.e., a subjective test of prime visibility), or both.	k = 74 Objective k = 57 Subjective k = 10 Both k = 7	k = 38 Objective k = 27 Subjective k = 9 Both k = 2
d'	Continuous	A measure often used to test prime visibility is the d' measure, a sensitivity measure based on signal detection theory. Participants receive the same stimuli and are asked to perform the same task as before but now on the primes instead of the targets. A d' value of 0 reflects chance-level visibility of the primes.	k = 58 M = 0.19, SD = 0.20 Range = −0.06 to 0.66	k = 4 M = 0.05, SD = 0.04 Range = −0.01 to 0.08

Note. k = number of effect sizes in the category.

Appendix B

Effect sizes ($\hat{\theta}_{jk}$), Standard Errors (SE_{jk}), and the Study Characteristics for the Semantic Categorization Conditions

Study	N	Pop	Prime format	Prime novel	Target format	Cat size	Set size	Target rep	Trials	P–T format	Prime dur	SOA	Masking	Visib meas	Obj/subj	d'	$\hat{\theta}_{jk}$	SE_{jk}
Bodner & Dypvik (2005)																		
Condition a	20	1	2	3	2	1	6	60	360	1	50	50	2	1	2		1.17	0.32
Condition b	20	1	2	3	2	1	6	60	360	1	50	50	2	1	2		0.13	0.24
Condition c	20	1	1	3	1	1	6	60	360	1	42	42	2	1	2		1.75	0.41
Condition d	20	1	1	3	1	1	6	60	360	1	42	42	2	1	2		0.42	0.25
Condition e	23	1	1	1	1	1	8	40	320	1	42	42	2	1	2		1.14	0.29
Condition f	24	1	1	1	1	1	8	40	320	1	42	42	2	1	2		0.37	0.22
Condition g	44	1	1	1	1	1	8	40	320	1	42	42	2	1	2		1.33	0.22
Condition h	57	1	1	1	1	1	8	40	320	1	42	42	2	1	2		1.00	0.17
Condition i	20	1	1	1	1	1	8	20	160	1	42	42	2	1	3		1.53	0.37
Condition j	20	1	1	1	1	1	8	20	160	1	42	42	2	1	3		1.64	0.39
Bueno & Frenck-Mestre (2002)	48	1	2	2	2	2	40	1	40	1	57	71	1	1	1		0.43	0.15
Damian (2001)																		
Condition a	16	1	2	1	2	2	12	10	120	1	43	72	1	1	1	.06	0.62	0.30
Condition b	16	1	2	1	2	2	12	10	120	1	43	72	1	1	1	.05	0.68	0.31
Condition c	16	1	2	2	2	2	12	10	120	1	43	72	1	1	1	.08	−0.15	0.27
Condition d	16	1	2	2	2	2	12	10	120	1	43	72	1	1	1	.02	0.00	0.27
Condition e	16	1	2	1	2	2	12	10	120	1	43	72	1	1	1	.12	0.60	0.30
Dehaene et al. (1998)																		
Condition a	12	2	3	1	3	1	8	8	64	1	43	114	1	1	1		1.76	0.59
Condition b	9[§]	2	3	1	3	1	8	8	64	1	43	114	1	0			1.31	0.61
Dell'Acqua & Grainger (1999)																		
Condition a	18	1	1	2	2	2	84	3	252	2	17	334	1	1	1	.05	0.53	0.27
Condition b	18	1	1	2	2	2	84	3	252	2	17	334	1	1	1	.05	0.58	0.28

(*table continues*)

Table (*continued*)

Study	N	Pop	Prime format	Prime novel	Target format	Cat size	Set size	Target rep	Trials	P–T format	Prime dur	SOA	Masking	Visib meas	Obj/subj	d'	$\hat{\theta}_{jk}$	SE_{jk}
Forster (2004)	22	1	2	2	2	1	70	2	140	1	55	55	2	0			1.62	
Forster et al. (2003)																		
Condition a	54	1	2	2	2	2	90	1	90	1	55	55	2	0			0.03	0.36
Condition b	54§	1	2	2	2	2	90	1	90	1	55	55	2	0			0.48	0.14
Condition c	54	1	2	2	2	2	90	1	90	1	41	41	2	0			0.08	0.15
Greenwald et al. (1989)	20	1	2	2	2	2	72	1	72	1	10	500	1	1	1		1.23	0.33
Kiesel et al. (2006)																		
Condition a	11	2	2	1	2	2	40	14	560	1	43	115	1	1	1	.17	0.36	0.35
Condition b	11§	2	2	2	2	2	40	14	560	1	43	115	1	1	1	.17	0.52	0.37
Condition c	11§	2	2	2	2	2	40	14	560	1	43	115	1	1	1	.17	0.78	0.41
Condition d	12	2	2	1	2	2	4	140	560	1	43	115	1	1	1	.24	1.02	0.43
Condition e	12§	2	2	2	2	2	4	140	560	1	43	115	1	1	1	.24	0.28	0.33
Condition f	12§	2	2	2	2	2	4	140	560	1	43	115	1	1	1	.24	−0.14	0.32
Kinoshita et al. (2006)																		
Condition a*	24	1	1	1	1	1	8	30	240	1	53	53	2	0			1.22	0.30
Condition b*	28	1	2	1	2	1	8	30	240	1	53	53	2	0			1.45	0.30
Koechlin et al. (1999)																		
Condition a	25	1	3	1	3	1	8	72	576	1	66	132	1	1	2		1.53	0.33
Condition b	24	1	1	1	1	1	8	72	576	1	66	132	1	1	2		1.04	0.28
Kunde et al. (2003)																		
Condition a	12	2	3	1	3	1	8	192	1536	1	43	115	1	1	1	.29	1.43	0.51
Condition b	12§	2	3	2	3	1	8	192	1536	1	43	115	1	1	1	.29	1.02	0.43
Condition c	12	2	1	1	1	1	4	192	768	1	29	101	1	1	1	.33	1.13	0.45
Condition d	12§	2	1	2	1	1	4	192	768	1	29	101	1	1	1	.33	0.48	0.35
Condition e	24	2	3	1	3	1	4	160	640	1	43	115	1	1	1	.22	0.82	0.25
Condition f	24§	2	3	2	3	1	4	160	640	1	43	115	1	1	1	.22	0.55	0.23
Kunde et al. (2005)																		
Condition a	16	2	3	1	3	1	4	160	640	2	72	144	1	1	1	.66	1.92	0.50
Condition b	16	2	3	2	3	1	4	160	640	2	72	144	1	1	1	.66	2.06	0.53
Naccache & Dehaene (2001)																		
Condition a	18	1	3	1	3	1	8	64	512	1	43	114	1	1	3	.60	0.88	0.31
Condition b	18§	1	3	2	3	1	8	64	512	1	43	114	1	1	3	.60	0.70	0.29
Condition c	18	2	3	1	1	1	8	64	512	2	43	114	1	1	3	.01	0.98	0.32
Condition d	18§	2	3	2	1	1	8	64	512	2	43	114	1	1	3	.01	0.76	0.29
Pohl et al. (2004)																		
Condition a*	18	1	2	1	2	2	4	140	560	1	43	115	1	1	1	.08	0.00	0.25
Condition b*	18§	1	2	2	2	2	4	140	560	1	43	115	1	1	1	.08	0.19	0.25
Condition c*	18§	1	2	2	2	2	4	140	560	1	43	115	1	1	1	.08	−0.12	0.25
Condition d*	17	1	2	1	2	2	40	14	560	1	43	115	1	1	1	.07	0.70	0.3
Condition e*	17§	1	2	2	2	2	40	14	560	1	43	115	1	1	1	.07	0.04	0.26
Condition f*	16	1	2	1	2	2	40	14	560	1	43	115	1	1	1	.05	0.83	0.33
Condition g*	16§	1	2	2	2	2	40	14	560	1	43	115	1	1	1	.05	−0.32	0.28
Condition h*	16§	1	2	2	2	2	40	14	560	1	43	115	1	1	1	.05	−0.11	0.27
Pohl et al. (2005)																		
Condition a*	11	1	2	1	2	2	40	20	800	1	43	115	1	1	1	.05	−0.12	0.34
Condition b*	11§	1	2	2	2	2	40	20	800	1	43	115	1	1	1	.05	0.00	0.34
Condition c*	11§	1	2	2	2	2	40	20	800	1	43	115	1	1	1	.05	0.15	0.34
Condition d*	11§	1	2	2	2	2	40	20	800	1	43	115	1	1	1	.05	0.60	0.38
Condition e*	11§	1	2	2	2	2	40	20	800	1	43	115	1	1	1	.05	−0.31	0.35
Pohl et al. (2006)																		
Condition a*	12	1	1	1	1	2	4	160	640	1	28	100	1	1	1	.35	1.28	0.48
Condition b*	12§	1	1	2	1	2	4	160	640	1	28	100	1	1	1	.35	0.07	0.32
Condition c*	12§	1	1	2	1	2	4	160	640	1	28	100	1	1	1	.35	0.39	0.34
Condition d*	12§	1	1	2	1	2	4	160	640	1	28	100	1	1	1	.35	−0.11	0.32
Condition e*	12	1	1	1	1	2	40	16	640	1	28	100	1	1	1	.08	0.95	(0.41)
Condition f*	12§	1	1	2	1	2	40	16	640	1	28	100	1	1	1	.08	1.05	(0.43)
Condition g*	12§	1	1	2	1	2	40	16	640	1	28	100	1	1	1	.08	0.66	(0.37)
Condition h*	12§	1	1	2	1	2	40	16	640	1	28	100	1	1	1	.08	−0.20	(0.32)
Condition i*	12	1	1	1	1	2	40	16	640	1	28	100	1	1	1	.64	1.88	0.61
Condition j*	12§	1	1	2	1	2	40	16	640	1	28	100	1	1	1	.64	0.75	0.38
Reynvoet et al. (2002)																		
Condition a	16	1	1	1	1	1	6	108	648	1	57	114	1	0			1.64	0.45
Condition b	16§	1	1	2	1	1	6	108	648	1	57	114	1	0			1.35	0.40

(*table continues*)

Table (*continued*)

Study	N	Pop	Prime format	Prime novel	Target format	Cat size	Set size	Target rep	Trials	P–T format	Prime dur	SOA	Masking	Visib meas	Obj/subj	d'	$\hat{\theta}_{jk}$	SE_{jk}
Condition c	16	1	1	1	1	1	6	80	480	1	57	114	1	0			1.51	0.43
Condition d	16§	1	1	2	1	1	6	80	480	1	57	114	1	0			1.03	0.35
Condition e	16	1	2	1	1	1	6	80	480	2	57	114	1	0			1.72	0.47
Condition f	16§	1	2	2	1	1	6	80	480	2	57	114	1	0			0.62	0.30
Rouibah et al. (1999)	80	1	2	2	2	2	60	4	240	1	49	273	1	0			0.82	0.13
Théoret et al. (2004)	6	2	3	1	3	1	16	21	336	1	43	114	1	1	3		1.55	1.03
Van den Bussche & Reynvoet (2006)																		
Condition a*	24	1	1	1	1	1	4	80	320	1	33	50	1	1	1	.46	1.37	0.31
Condition b*	24§	1	1	2	1	1	4	80	320	1	33	50	1	1	1	.46	1.10	0.28
Condition c*	24	1	2	1	2	2	12	12	144	1	33	50	1	1	1	−.06	1.31	0.31
Condition d*	24§	1	2	2	2	2	12	12	144	1	33	50	1	1	1	−.06	0.25	0.22
Van Opstal et al. (2005a)																		
Condition a	23	2	3	1	3	1	4	160	640	1	33	100	1	1	1	.00	0.77	0.26
Condition b	23§	2	3	2	3	1	4	160	640	1	33	100	1	1	1	.00	0.27	0.22
Condition c	23	2	3	1	3	1	4	160	640	1	33	100	1	1	1	.15	1.80	0.38
Condition d	23§	2	3	2	3	1	4	160	640	1	33	100	1	1	1	.15	1.76	0.38
Van Opstal et al. (2005b)																		
Condition a	16	2	1	1	1	1	4	192	768	1	33	100	1	1	1	.18	0.73	0.31
Condition b	16§	2	1	1	1	1	4	192	768	1	33	100	1	1	1	.18	1.11	0.37

Note. Effect sizes for condition j within study $k(\hat{\theta}_{jk})$ that are positive indicate positive priming effects. N = sample size; Pop = population; Prime novel = prime novelty; Cat size = category size; Target rep = target repetitions; P–T format = prime–target format; Prime dur = prime duration; SOA = stimulus onset asynchrony; Visib meas = visibility measured; Obj/subj = nature of visibility measure; d' = objective measure of prime visibility. Population: 1 = academics, 2 = nonacademics. Primes: 1 = symbols, 2 = words, 3 = both. Prime novelty: 1 = repeated primes, 2 = novel primes; 3 = both. Targets: 1 = symbols, 2 = words; 3 = both. Category size: 1 = small, 2 = large. Prime–target format: 1 = same, 2 = different. Masking: 1 = backward pattern mask, 2 = backward target mask. Visibility measured: 0 = no, 1 = yes. Obj/subj: 1 = objective, 2 = subjective, 3 = both.
* Unpublished data.
§ (Part of the) same sample used as in a previous condition of the same study.

Appendix C

Effect sizes ($\hat{\theta}_{jk}$), Standard Errors (SE_{jk}), and the Study Characteristics for the Lexical Decision and Naming Conditions

Study	N	Pop	Task	Prime format	Prime novel	Target format	Cat size	Set size	Target rep	Trials	P–T format	Prime dur	SOA	Masking	Visib meas	Obj/subj	d'	$\hat{\theta}_{jk}$	SE_{jk}
Alameda et al. (2003)																			
Condition a	50	1	2	2	2	1	2	50	1	50	2	84	98	1	0			0.89	0.17
Condition b	50	1	2	2	2	1	2	50	1	50	2	57	71	1	0			0.34	0.15
Condition c	50	1	1	2	2	1	2	50	1	100	2	84	98	1	0			0.70	0.16
Bajo et al. (2003)																			
Condition a	30	1	2	2	2	1	2	30	1	30	2	50	64	1	0			0.28	0.19
Condition b	30§	1	2	2	2	1	2	30	1	30	2	75	89	1	0			−0.07	0.19
Bodner & Masson (2003)																			
Condition a	100	1	1	2	2	2	2	200	1	400	1	45	45	2	1	2		0.76	0.11
Condition b	50§	1	1	2	2	2	2	200	1	400	1	45	45	2	1	1		0.94	0.18
Condition c	40	1	1	2	2	2	2	200	1	400	1	45	45	2	1	2		0.69	0.18
Bourassa & Besner (1998)																			
Condition a	130	1	1	2	2	2	2	80	1	160	1	40	80	1	1	2		0.37	0.09
Condition b	132	1	1	2	2	2	2	80	1	160	1	40	340	1	1	2		0.33	0.09
Brown & Besner (2002)	72	1	1	2	2	2	2	96	1	192	1	34	784	1	0			0.35	0.12
Carr & Dagenbach (1990)	12	1	1	2	2	2	2	72	1	144	1	10		1	1	1		1.08	0.44
Dell'Acqua & Grainger (1999)	19§	1	2	1	2	1	2	72	3	216	1	17	334	1	1	1	0.07	0.71	0.28
Devlin et al. (2004)	11	2	1	2	2	2	2	56	1	168	1	33	33	2	1	1		0.45	0.36
Devlin et al. (2006)	11	2	1	2	2	2	2	56	1	280	1	33	33	2	1	1		0.45	0.36
Finkbeiner & Caramazza (2006)																			
Condition a	18	1	2	2	2	1	2	46	2	92	2	53	53	2	1	2		0.57	0.28
Condition b	20	1	2	2	2	1	2	38	2	76	2	53	66	1	1	2		0.47	0.25

(*table continues*)

Table (*continued*)

Study	N	Pop	Task	Prime format	Prime novel	Target format	Cat size	Set size	Target rep	Trials	P–T format	Prime dur	SOA	Masking	Visib meas	Obj/subj	d'	$\hat{\theta}_{jk}$	SE_{jk}
Forster & Hector (2005)																			
Condition a[*]	20	1	1	2	2	2					1	42		1	1	1		0.48	0.25
Condition b[*]	20[§]	1	1	2	2	2					1	42		1	1	1		0.22	0.24
Condition c[*]	18	1	1	2	2	2					1	55		1	1	1		0.27	0.26
Condition d[*]	18[§]	1	1	2	2	2					1	55		1	1	1		0.53	0.27
Condition e[*]	20	1	1	2	2	2					1	69		1	1	1		0.22	0.24
Condition f[*]	20[§]	1	1	2	2	2					1	69		1	1	1		0.43	0.25
Frost et al. (1997)	48	2	1	2	2	2	2	48	1	96	1	43	43	2	0			−0.10	0.15
Grossi (2006)																			
Condition a	20	1	1	2	2	2	2	200	1	400	1	50	50	2	1	1	0.07	0.81	0.28
Condition b	20	1	1	2	2	2	2	200	1	400	1	50	50	2	1	1	0.08	0.59	0.26
Hines et al. (1986)																			
Condition a	30	1	2	1	2	2	2	30	4	120	2	42	642	1	1	1		0.40	0.20
Condition b	57	1	2	1	2	2	2	36	2	72	2	19	499	1	1	1		0.15	0.14
Hines (1993)	80	1	2	1	2	2	2	32	2	64	2	17	497	1	1	1		0.21	0.11
Kamphuis et al. (2005)																			
Condition a	20	1	1	2	1	2	1	18	4	144	1	40	67	1	0			−0.24	0.24
Condition b	20[§]	1	1	2	1	2	1	18	4	144	1	40	67	1	0			−0.13	0.24
Kemp-Wheeler & Hill (1988)																			
Condition a	9	1	1	2	2	2	2	20	1	40	1	51	551	1	1	1		1.39	0.63
Condition b	10	1	1	2	2	2	2	20	1	40	1	37	537	1	1	1		1.29	0.55
Condition c	9	1	1	2	2	2	2	20	1	40	1	24	524	1	1	1		1.75	0.74
Condition d	9	1	1	2	2	2	2	20	1	40	1	16	516	1	1	1		1.85	0.77
Kiefer (2002)	24	2	1	2	2	2	2	320	1	640	1	33	67	1	1	3	−0.01	1.09	0.28
Kiefer & Brendel (2006)																			
Condition a	16	2	1	2	2	2	2	320	1	800	1	33	67	1	1	1		0.96	0.34
Condition b	16[§]	2	1	2	2	2	2	320	1	800	1	33	200	1	1	1		0.46	0.29
Condition c	16	2	1	2	2	2	2	320	1	800	1	33	67	1	1	1		0.97	0.34
Condition d	16[§]	2	1	2	2	2	2	320	1	800	1	33	200	1	1	1		1.10	0.36
Kiefer & Spitzer (2000)																			
Condition a	20	2	1	2	2	2	2	320	1	640	1	50	67	1	1	1		1.00	0.30
Condition b	20[§]	2	1	2	2	2	2	320	1	640	1	50	200	1	1	1		1.07	0.31
Marcel (1983)	12	1	1	2	2	2	2	90	1	180	1	10		1	1	1		1.34	0.49
McBride et al. (2003)[*]	20	2	1	2	2	2					1	67		1	0			0.49	0.25
O'Seaghdha & Marin (1997)	32	1	2	2	2	2	2	60	1	120	1	57	57	2	0			0.51	0.20
Perea & Gotor (1997)																			
Condition a	22	1	1	2	2	2	2	64	1	128	1	33	33	2	0			0.27	0.23
Condition b	22	1	1	2	2	2	2	64	1	128	1	50	50	2	0			0.93	0.28
Condition c	22	1	1	2	2	2	2	64	1	128	1	67	67	2	0			0.46	0.24
Condition d	22	1	2	2	2	2	2	64	1	64	1	33	33	2	0			0.00	0.22
Condition e	22	1	2	2	2	2	2	64	1	64	1	50	50	2	0			0.38	0.23
Condition f	22	1	2	2	2	2	2	64	1	64	1	67	67	2	0			0.58	0.25
Condition g	26	1	1	2	2	2	2	64	1	128	1	67	67	2	0			0.55	0.22
Condition h	26	1	2	2	2	2	2	64	1	64	1	67	67	2	0			0.45	0.22
Perea & Lupker (2003)																			
Condition a	120	1	1	2	2	2	2	120	1	240	1	40	80	1	0			0.31	0.09
Condition b	36	1	1	2	2	2	2	120	1	240	1	40	80	1	0			1.03	0.22
Perea & Rosa (2002a)																			
Condition a	48	1	1	2	2	2	2	24	1	132	1	66	66	2	0			0.31	0.15
Condition b	48	1	2	2	2	2	2	24	1	66	1	66	66	2	0			0.32	0.15
Perea & Rosa (2002b)																			
Condition a	32	1	1	2	2	2	2	66	1	132	1	66	66	2	0			0.39	0.19
Condition b	32	1	1	2	2	2	2	66	1	132	1	83	83	2	0			0.19	0.45
Condition c	32	1	1	2	2	2	2	66	1	132	1	66	66	2	0			0.02	0.18
Condition d	32	1	1	2	2	2	2	66	1	132	1	83	83	2	0			0.85	0.22
Ruz et al. (2003)	45	1	1	2	2	2	2	45	8	360	1	13		1	1	3		0.31	0.16
Sassenberg & Moskowitz (2005)	15	1	1	2	2	2	2	6	4	48	1	50	50	2	0			0.79	0.33
Sebba & Forster (1989)[*]	45	1	1	2	2	2					1	50		1	0			−0.15	0.15

(*table continues*)

Table (*continued*)

Study	N	Pop	Task	Prime format	Prime novel	Target format	Cat size	Set size	Target rep	Trials	P–T format	Prime dur	SOA	Masking	Visib meas	Obj/subj	d'	$\hat{\theta}_{jk}$	SE_{jk}
Williams (1996)																			
Condition a	24	1	1	2	2	2	2	22	1	44	1	50	50	2	1	2		0.59	0.24
Condition b	16	1	1	2	2	2	2	20	1	40	1	50	50	2	1	2		0.61	0.30
Condition c	26	1	1	2	2	2	2	40	1	80	1	50	50	2	1	2		0.72	0.23
Zhou et al. (1999)	29	1	1	1	2	1	2	45	1	125	1	57	57	2	0			0.41	0.20

Note. Effect sizes for condition j within study k (θ_{jk}) that are positive indicate positive priming effects. N = sample size; Pop = population; Prime novel = prime novelty; Cat size = category size; Target rep = target repetitions; P–T format = prime–target format; Prime dur = prime duration; SOA = stimulus onset asynchrony; Visib meas = visibility measured; Obj/subj = nature of visibility measure; d' = objective measure of prime visibility. Population: 1 = academics, 2 = nonacademics. Primes: 1 = symbols, 2 = words. Task: 1 = lexical decision, 2 = naming. Prime novelty: 1 = repeated primes, 2 = novel primes. Targets: 1 = symbols, 2 = words. Category size: 1 = small, 2 = large. Prime–target format: 1 = same, 2 = different. Masking: 1 = backward pattern mask, 2 = backward target mask. Visibility measured: 0 = no, 1 = yes. Obj/subj: 1 = objective, 2 = subjective, 3 = both.
* Unpublished data.
§ (Part of the) same sample used as in a previous condition of the same study.

Received April 14, 2008
Revision received November 17, 2008
Accepted November 26, 2008 ■