

Masked Priming Effects in Semantic Categorization Are Independent of Category Size

Eva Van den Bussche and Bert Reynvoet

Department of Psychology, University of Leuven, Belgium

Abstract. The question whether subliminal primes can activate their semantic meaning or not is still of interest today. Three different competing theories have tried to account for the often inconsistent research results: The semantic categorization hypothesis, the direct motor specification hypothesis, and the category search model. The present study aimed to shed light on these different points of view by examining the role of category size in response congruency effects when novel primes are used. Three experiments were conducted and a transparent pattern of results emerged: Significant priming effects were obtained across different tasks, irrespective of category size and irrespective of stimulus set size. The findings are discussed in terms of the three theoretical frameworks. It becomes clear that the present results provide strong evidence in favor of the semantic categorization hypothesis, which assumes semantic processing of subliminal primes.

Keywords: masked priming, category size, semantic categorization, congruency effect

Introduction

The research history of subliminal priming reflects a long-standing debate in cognitive psychology, which is still very much alive today. Marcel (1983) found that unconsciously presented words significantly primed semantic associates. Although his results were described as “startling” and “counterintuitive” (Fowler, Wolford, Slade, & Tassinari, 1981), successful replications accumulated throughout the years (e.g., Balota, 1983; Fowler et al., 1981; Forster & Davis, 1984; Greenwald, Klinger, & Liu, 1989). As the supporting evidence grew, so did the skepticism. Holender (1986) reviewed the then-available support of masked priming as contaminated by a variety of problems such as lack of reliability and poor assessment of whether stimuli were actually unconsciously presented. During the last decade, methodology and measurement strategy have advanced, producing more reliable research (e.g., Draine & Greenwald, 1998; Greenwald, Draine, & Abrams, 1996; Merikle, Smilek, & Eastwood, 2001). This new and improved research flow provided strong evidence for reproducible subliminal priming effects. These evolutions led to a shift in the focus of the debate. As the existence of subliminal priming effects was now no longer questioned, and regarded as well-established and more trustworthy, the debate progressed beyond existence claims (Greenwald, Abrams, Naccache, & Dehaene, 2003). Currently, one of the most puzzling and central questions is whether subliminal primes can activate their semantic meaning, or not.

To assess the depth of processing of subliminal numerical primes, Dehaene and colleagues (Dehaene et al., 1998) used a simple semantic categorization task where subjects needed to categorize numbers between 1 and 9 as smaller or larger than 5. Two types of trials could be distinguished: *Congruent* trials where prime and target evoked the same response and *incongruent* trials where prime and target evoked opposite responses. They found that the congruent trials were responded to faster than incongruent trials. Furthermore, this significant response congruency effect (RCE) was independent of notation format (Arabic digits vs. written number words) and did not reflect a word identity effect. Using brain-imaging techniques, they discovered that when subjects participated in an overt semantic categorization task with visible targets, they also unconsciously applied the task instructions to the subliminal primes. They concluded that subliminal primes are processed in a series of processing stages, including semantic categorization (i.e., *semantic categorization hypothesis*).

The findings of Dehaene et al. (1998) were criticized by Damian (2001), who shed a different light on their findings. Although in a first experiment he was able to replicate and extend Dehaene et al.'s results using nouns referring to concrete objects, the RCE disappeared in a subsequent experiment where (contrary to Dehaene et al. and their own Experiment 1) different stimuli were used as primes and targets (see also Abrams & Greenwald, 2000, Experiment 3). These findings do not support the semantic categorization hypothesis: If subliminal primes are se-

mentally processed, the RCE should emerge irrespective of whether prime and target set are identical. Alternatively, these results were interpreted in terms of the *direct motor specification hypothesis* (Neumann & Klotz, 1994). Within this framework it is assumed that when the stimulus set is limited and prime and target set comprise the same stimuli, subjects can learn to unconsciously associate these stimuli with the adequate responses, bypassing semantic access, and resulting in automatized stimulus-response mappings (Damian, 2001). The failure to sustain the RCE when prime and target set consist of different stimuli is direct evidence in favor of this hypothesis.

While some authors failed to establish a congruency effect using novel primes, others succeeded. For instance, Naccache and Dehaene (2001) replicated the number priming experiment from Dehaene et al. (1998) using both "old set primes" (primes that also appeared as targets) and "new set primes" (novel, unseen primes that never appeared as targets). Their results clearly indicated that the RCE generalized to novel primes. These findings are difficult to reconcile with the direct motor specification hypothesis, which would predict that priming should be obtained with old set primes, but not with new set primes since stimulus-response mappings could not be established. On the other hand, if subjects unconsciously apply the task instructions to the primes and, thus, categorize the primes at a semantic level (as proposed by Dehaene et al., 1998), it seems logical that the RCE should extend to the novel primes. The results of Naccache and Dehaene (2001) clearly favor the latter interpretation of subliminal priming. This finding was later confirmed by Reynvoet and colleagues (Reynvoet, Caessens, & Brysbaert, 2002; Reynvoet, Gevers, & Caessens, 2005) and Greenwald et al. (2003), who also reported a reliable RCE when primes were not part of the target set.

Forster, Mohan, and Hector (2003, Experiment 3) attempted to clarify these contradictory priming effects in semantic categorization when novel primes are used. In their experiment, words needed to be categorized as animals or nonanimals. Primes never occurred as targets. In line with Damian (2001) and Abrams and Greenwald (2000) they found no congruency effect for the animal targets (exemplars). However, contrary to their expectations, the authors found a significant RCE of 15 ms for the nonanimals (nonexemplars). This pattern of results had been confirmed in a previous (identical) experiment of the same authors, with shorter prime duration (Forster et al., 2003, pp. 26–27). These findings imply that neither the semantic categorization hypothesis (that would predict priming to occur in *both* exemplars nor nonexemplars), nor the direct motor specification hypothesis (that would predict priming to occur in *neither* exemplars nor nonexemplars) present a complete picture of subliminal priming. The authors realized that additional processes must be at work and wondered whether the task conditions can introduce certain factors or circumstances that modify or mediate the priming effects. The asymmetry between exemplars (ani-

mals) and nonexemplars (nonanimals) indicates that the type of category (numbers, animals, nonanimals, etc.) chosen in the experimental design might be another possible side condition mediating subliminal priming effects.

To further explore the role of the categories used in semantic categorization experiments, Forster (2004, Experiment 4) investigated the effects of *category size* on subliminal priming effects. Contrary to Forster's previous study where he used a broad category (animals), he now used small categories (months, numbers, farm animals, type of dog, and body parts) to directly assess the impact of the category size on priming effects. Again, novel primes were used. Subjects were presented with a category and had to decide whether the presented targets were a member of this category or not. Contrary to his previous findings, a substantial congruency effect was obtained for both exemplars and nonexemplars, leading Forster (2004) to conclude that there is a clear contrast between subliminal priming effects for small categories (e.g., numbers) and broad categories (e.g., animals). The divergent results are explained in terms of a *category search model* (Landauer & Freedman, 1968). This model states that decisions about category membership (e.g., animal or not?) are made by a search of the category members. From this point of view, it directly follows that response times (RTs) for "Yes" (e.g., animal) and "No" (e.g., nonanimal) decisions are influenced by the category size and that a "No" decision will require an exhaustive search of the category. When a small category is used, the probability that an exhaustive search can be completed is high so that an implicit "Yes" or "No" decision about the category membership of the prime can be reached before an explicit decision about the target. If the decision based on the target matches the earlier implicit decision about the prime (congruent trials), no problem arises and reaction times will be unaffected, but if the decisions differ (incongruent trials), RTs are slowed down, resulting in a congruency effect. In contrast, when using a broad category, an exhaustive search is impossible. On some occasions, an exemplar prime can be classified (a "Yes" decision does not always require an exhaustive search), but nonexemplars can never be classified (a "No" decision needs an exhaustive search). Reasoning along these lines, Forster (2004) concluded that nonexemplar primes will never interfere with the explicit decision made on the targets and congruency effects will be absent. Exemplar primes have the potential to, at least on some occasions, interfere with the explicit decision on nonexemplar targets leading to a small congruency effect on nonexemplar targets.

While such an explanation in terms of category size can neatly explain the results obtained by Forster and colleagues (2003, 2004), it cannot account for all findings concerning subliminal priming. For instance, Dell'Acqua and Grainger (1999, Experiment 1) examined the effect of unconsciously presented picture primes on word targets in a semantic categorization task. Subjects needed to categorize the targets as natural things or artifacts. Clearly, broad

categories were used and no congruency effects would be expected following Forster's line of reasoning. However, the results showed a significant RCE, with incongruent trials (where prime picture and target word belonged to different categories) being responded to slower than trials where prime picture and target word belonged to the same category. Furthermore, Reynvoet et al. (2005) showed that the RCE using novel primes is not limited to a small set of numerical stimuli, but also extends to a somewhat larger set, such as letter stimuli.

Reviewing these contradictory research findings and proposed theories in semantic categorization, the need to clarify the influence of category size on response congruency effects using novel primes becomes clear. First, clarifying this issue will shed light on previous equivocal research findings. Second, and most importantly, resolving this issue will have implications for the theoretical frameworks proposed. If RCEs of novel primes indeed occur irrespective of category size, this would provide further evidence in favor of the semantic categorization theory (Dehaene et al., 1998; Naccache & Dehaene, 2001; Reynvoet et al., 2002, 2005) and against the direct motor specification hypothesis (Damian, 2001; Neumann & Klotz, 1994). Furthermore, it would contradict the category search model proposed by Forster (2004).

In conclusion, the aim of this study is straightforward: to seek such clarification by using both small and broad categories and, by doing this, solving one piece in the puzzle as to what extent subliminally presented stimuli are semantically processed.

Experiment 1

Experiment 1 consisted of three separate parts. In the first part (Experiment 1a) numbers needed to be categorized as smaller or larger than 5. Thus, a small category was used and a very central meaning (magnitude) of this category had to be evaluated. In the second part (Experiment 1b) body parts needed to be categorized as above or below the pelvis. Again, a small category was used but the stimuli were no longer numbers, which some authors have claimed may form a special case (Damian, 2001; Greenwald et al., 2003). Finally, in the third part (Experiment 1c) a broad category had to be judged: Animals needed to be categorized as smaller or larger than a reference animal, namely a dog of the type Labrador. Subjects categorized all stimuli by pressing one of two buttons on the keyboard. The targets were preceded by primes that were never presented as targets. This design allowed us to investigate whether the RCE is present for both narrow and broad categories using novel primes. We carefully controlled for prime visibility by using objective measures of prime visibility to ensure that primes could not be consciously reported.

Method

Participants

Sixteen first year Educational Sciences students participated in the experiment (all three parts) as partial fulfillment of a course requirement. One subject was omitted from further analyses because he was substantially (+ 2SD) slower than the other subjects. Thus, the final sample consisted of 15 subjects (2 male, 13 female) with a mean age of 19.53 (range 17–28).

Apparatus

Stimuli were presented on a 15-inch color screen connected to a computer running the Windows operating system. Stimulus delivery and the recording of behavioral data (reaction time and accuracy) were controlled by E-prime (www.pstnet.com; Psychology Software Tools).

Procedure

The procedure was the same for all three parts of this experiment. Each trial consisted of the following sequence. First, a forward mask was shown for 480 ms. The mask consisted of six hash marks (#####), 3.2 cm in width and 0.7 cm in height. Then a prime stimulus was presented for 33 ms and was immediately replaced by a backward mask (#####) for 17 ms. Next the target was presented and remained on the screen until the participant had responded. The intertrial interval was 1000 ms. All presentations were synchronized with the vertical refresh cycle of the screen (16.7 ms).

In each part of the experiment, the stimuli were divided into a prime set and a target set. Half of the number words were only presented as primes, whereas the other half were only presented as targets. All targets were presented in capital letters, while the primes were presented in lower-case letters. Primes and targets were presented as white letters on a black background and ranged from 0.9 cm to 3.3 cm in width and 0.7 cm in height. Half of the trials were congruent trials (both prime and target evoked the same response), whereas the other half were incongruent trials (prime and target evoked different responses). Participants were told that they would see word targets, which needed to be classified in two categories. The response assignment was varied across participants: half of the subjects had to respond with a left-hand key press ('a') for Category 1 and with a right-hand key press ('p') for Category 2; the other half did the experiment with the reversed response assignment. Participants were instructed to respond as quickly as possible and to avoid mistakes as far as possible. Crucially, all participants were unaware of the purpose of the study and were told that a series of neutral visual flashes was presented to indicate the target.

Stimuli

Experiment 1a

The stimuli consisted of eight Dutch number words. The numbers used were one, two, three, four, six, seven, eight, and nine. The number words two, three, seven, and eight were only presented as primes, whereas one, four, six, and nine were only presented as targets. In both the prime and the target group half of the stimuli were smaller than five and half were larger than five. All possible prime-target combinations were formed; in total this led to 16 prime-target combinations (four primes \times four targets). Each participant started with a training block of 16 trials in which each target was presented four times, without a prime. Then the experimental session started containing 10 blocks of 16 trials (cf. the 16 prime-target combinations). Participants were instructed to classify the number word targets as smaller or larger than five.

Experiment 1b

For the second part of the experiment another small category was used, namely body parts. The stimuli consisted of 16 body parts words each of which consisted of no more than five letters. Half of the body parts words were located below the pelvis, and half were located above the pelvis. All stimuli were medium to high frequency words (log frequencies ranging from 0.90 to 3.01) and both classes (i.e., below and above pelvis) were matched on frequency using the WordGen program of Duyck, Desmet, Verbeke, and Brysbaert (2004). Half of the stimuli were only presented as primes, whereas the other half were only presented as targets. All possible prime-target combinations were formed; in total this led to 64 prime-target combinations (eight primes \times eight targets). Each participant started with a training block of 16 trials in which each target was presented twice, without a prime. Then the experimental session started containing three blocks of 64 trials (cf. the 64 prime-target combinations). Participants were told to classify the body parts word targets as below or above the pelvis.

Experiment 1c

For this third part of the experiment a broad category was used, namely animals. The stimuli consisted of 16 animal words each of which consisted of no more than five letters. In both the prime and the target group half of the stimuli were smaller than a reference animal, a dog of the type Labrador, and half were larger than a Labrador. All stimuli were medium to high frequency words (log frequencies ranging from 0.60 to 2.20) and both classes (i.e., smaller and larger than a Labrador) were matched on frequency using the WordGen program of Duyck et al. (2004). Half of the animal words (eight) were only presented as primes, whereas the other half (eight) were only presented as targets. All possible prime-target combinations were formed;

in total this led to 64 prime-target combinations (eight primes \times eight targets). As far as possible, the prime-target pairs were matched on log frequency. Each participant started with a training block of 16 trials in which each target was presented twice, without a prime. Then the experimental session started containing three blocks of 64 trials (cf. the 64 prime-target combinations). Participants had to classify the animal word targets as smaller or larger than a Labrador, which is widely regarded as a highly prototypical dog race in Belgium. Included in the instructions was a picture of a Labrador to ensure that the reference animal was clear to the participants.

All stimuli and pairings used in Experiment 1 are listed in Appendix A.

Prime Visibility

Prime visibility was assessed for each participant using objective visibility tests. Immediately after the end of the last part of the experiment, the participants were informed about the sequence of a priming trial. They were asked to participate in three posttests, corresponding to each of the three parts of the experiment. Each posttest consisted of 64 trials: For Experiment 1a all prime-target combinations were presented four times (4×16 trials); for Experiment 1b and 1c all prime-target combinations were presented once (1×64 trials). The participants were instructed to apply the same instructions they received before (cf. Experiments 1a, 1b and 1c) to the prime instead of the target. If they were unable to categorize the prime, they were forced to guess (forced choice). Presentation conditions were the same as during the experiments.

Results

RT and Error Rate Analysis

Wrong responses (on average, 6.1% of the data) were discarded. For each of the three parts of this experiment, median RTs from correct responses were then submitted to a repeated measures analysis with congruency (congruent or incongruent) as within-subject factor. Median RTs and mean error rates as a function of congruency for Experiments 1a, 1b, and 1c are listed in Table 1.

Experiment 1a

For the number task, we found an effect of response congruency, $F(1, 14) = 59.66, p < .001$, with congruent trials being responded to 20 ms faster than incongruent trials. The same repeated measures analysis performed on error rates revealed a significant effect of congruency, $F(1, 14) = 7.40, p < .05$. More errors were made on incongruent trials compared to congruent trials (6.8% vs. 3.8%).

Table 1. Median RTs (*SD*) and mean percent error rates (*SD*) as a function of congruency for Experiments 1a, 1b, and 1c, and the differences (RCEs) between congruent and incongruent trials (incongruent – congruent)

	Median RTs (<i>SD</i>)			Error rates (<i>SD</i>)		
	Congruent	Incongruent	Difference	Congruent	Incongruent	Difference
Experiment 1a	468.9 (21.6)	488.9 (21.9)	20.0	3.8 (2.6)	6.8 (3.9)	3.0
Experiment 1b	553.0 (43.6)	560.6 (47.3)	7.6	5.9 (3.1)	9.4 (6.0)	3.5
Experiment 1c	487.5 (28.8)	512.1 (19.2)	24.6	3.7 (2.9)	6.3 (3.4)	2.6

Note. Experiment 1a: categorization of number word targets as smaller or larger than 5; Experiment 1b: categorization of body parts word targets as below or above the pelvis; Experiment 1c: categorization of animal word targets as smaller or larger than a Labrador.

Experiment 1b

For the body parts task, we again found an effect of response congruency, $F(1, 14) = 5.78, p < .05$, with congruent trials being responded to 8 ms faster than incongruent trials. The same repeated measures analysis performed on error rates revealed a significant effect of congruency, $F(1, 14) = 7.96, p < .05$. More errors were made on incongruent trials compared to congruent trials (9.4% vs. 5.9%).

Experiment 1c

For the animal task, we also found an effect of response congruency, $F(1, 14) = 37.15, p < .001$, with congruent trials being responded to 25 ms faster than incongruent trials. The same repeated measures analysis performed on error rates revealed a significant effect of congruency, $F(1, 14) = 17.17, p < .01$. More errors were made on incongruent trials compared to congruent trials (6.3% vs. 3.7%).

Prime Visibility

Analyses of the posttests revealed that although the participants were fully informed about the object of the exercise, on average they could only classify 49% (for the number task), 49% (for the body parts task), and 51% (for the animal task) of the primes correctly. A direct measure of prime visibility (d') was calculated for each subject. The measures were obtained by treating one level of the response category (i.e., smaller than five, below the pelvis, smaller than a Labrador) as signal and the other level (i.e., larger than five, above the pelvis, larger than a Labrador) as noise. Mean d' values were, respectively, -0.029 , -0.087 , and 0.076 . Three t -tests against the null mean indicated that the mean observed d' values were not significantly different from 0, respectively, $t(14) = -0.28, p > .75$; $t(14) = -0.57, p > .55$; $t(14) = 0.57, p > .55$, indicating that the parameters were sufficient to guarantee unconscious presentation of primes. Using linear regression (e.g., Greenwald et al., 1996), nonsignificant correlations were found between the individual d' measures for the three tasks (numbers, body parts, and animals) and indices for the respective amounts of response priming, $r = .34, F(1, 13) = 1.69, p > .20$ for the number task; $r = -.11, F(1, 13) = 0.16, p > .65$ for the body parts task; $r = -.29, F(1, 13) = 0.72, p > .40$ for the

animal task. These results clearly indicate that the congruency effects were unrelated to the participants' perceptual sensitivity to the subliminal primes.

Discussion

In Experiment 1 we wanted to further explore the impact of category size on congruency priming effects when novel primes are used by using three different categories: numbers, body parts, and animals. We wanted to explore whether congruency effects are only obtained for small categories (cf. Forster, 2004) or can also be found using broad categories (cf. Dell'Acqua & Grainger, 1999). The present experiment consisted of three separate parts: Two parts used small categories (numbers and body parts) and one part used a broad category (animals). Because subjects completed all three parts of the experiment, congruency effects for each of these three categories could be easily compared. The results are clear-cut. First, we found that congruent number trials, congruent body parts trials, and congruent animal trials were responded to faster than their incongruent counterparts. Moreover, the RCE observed for the animal task (the broad category) was the most obvious. Secondly, significantly more errors were made on the incongruent trials compared to the congruent trials for all three categories. Again, this congruency effect was most obvious for the broad category.

Experiment 1 clearly extends congruency priming for novel primes beyond small categories. These results are in line with the findings of Dell'Acqua and Grainger (1999) who also found RCEs for broad categories. In contrast, the direct motor specification hypothesis proposed by Damian (2001) and the category search model proposed by Forster (2004) are not supported by these findings.

The present results lead to an interesting observation. In both our Experiment 1c and Experiment 3 of Damian (2001) a similar task was used: Subjects needed to judge the size of the targets relative to some reference, respectively, a dog and a rectangle. However, Damian did not obtain congruency effects when using novel primes. A possible, yet speculative, explanation for this discrepancy in results might lie in the nature of the stimuli. In our Experiment 1c, animals had to be judged on size, while Damian used stimuli from different categories like animals, fruits, and objects. It could be that "size" is a more central feature

for our animal stimuli compared to some of the stimuli used by Damian. A similar argument has been made for the comparison between the number domain and the word domain (Damian, 2001; Greenwald et al., 2003).

Two further comments can be made. First, one could argue that the task used for the broad (animal) category is hard to compare to the task employed by Forster et al. (2003). In our experiment, targets had to be categorized on the basis of one semantic feature (magnitude, locus on the body, size) of the stimulus class. Forster et al. (2003) used a wider task where subjects had to decide whether targets were animals or not, irrespective of the features they used to fulfill the task. Maybe the contradictory results originate from this difference in approach. Participants might have applied different strategies according to the task at hand. Second, since in Experiment 1 all stimulus sets were rather small, the limited set sizes could be responsible for the observed effects, which would be in line with the direct motor specification hypothesis. Damian (2001) cautioned against underestimating the power of repeated stimuli presentation, since undemanding tasks (like the tasks used in Experiment 1), automatize quickly. He recommended avoiding repeated exposure to the same stimuli. Therefore, two subsequent experiments were designed to ensure that possible different outcomes could not be attributed merely to a different experimental design or a limited stimulus set. In Experiment 2 subjects were instructed to classify words as animals or objects, without extracting a single feature. In Experiment 3, which replicates Forster's task, subjects were asked to categorize words as animals or nonanimals. Both experiments used a large stimulus set (104 stimuli).

Experiment 2

Method

Participants

Fourteen first year Educational Sciences students (10 female, 4 male) with a mean age of 23.36 (range 17–35) participated in the experiment as partial fulfillment of a course requirement.

Apparatus, Stimuli, and Procedure

Apparatus was the same as in Experiment 1. Two broad categories were used, namely animals and everyday objects. Targets needed to be categorized as being either animals or objects. The stimuli consisted of 52 animal words and 52 object words and each consisted of no more than five letters. All stimuli were medium to high frequency words (log frequencies ranging from 0.48 to 2.22 for the animal words and from 0.70 to 2.59 for the object words obtained by using the WordGen program of Duyck et al.,

2004). The stimuli were divided into four different groups: two served as prime sets and the other two as target sets. Half of the animal words (26) and half of the object words (26) were only presented as primes, whereas the other half of animal words (26) and object words (26) were only presented as targets.

All targets were presented twice in each block: They were primed either with a congruent prime (e.g., cat – APE, chair – PIANO), i.e., a prime that would lead to the same response as the target, or an incongruent prime (pot – APE, hawk – PIANO). As far as possible, the primes were of the same length as the targets. In total this led to 104 prime-target combinations (2 primes \times 52 targets). Each participant started with a training block of 52 trials in which each target was presented once, without a prime. Then the experimental session started containing three blocks of 104 trials (cf. the 104 prime-target combinations). The stimuli and pairings are listed in Appendix B. Participants were told that they would see word targets, which needed to be classified in two categories, animals or objects. Half of the trials were congruent trials (both prime and target evoked the same response), whereas the other half were incongruent trials (prime and target evoked different responses). As in Experiment 1, instructions varied between subjects. Participants were instructed to respond as quickly as possible and to avoid mistakes as far as possible.

Prime visibility was assessed analogous to Experiment 1. The posttest consisted of 104 trials: all prime-target combinations were presented once. The rest of the procedure was identical to Experiment 1.

Results

RT and Error Rate Analysis

Wrong responses (on average 4.8% of the data) were discarded. Median RTs from correct responses were entered in a repeated measures analysis with congruency (congruent or incongruent) and target class (animals or objects) as within-subject factors. Median RTs and mean error rates as a function of congruency and target class are listed in Table 2.

We found a significant main effect of congruency, $F(1, 13) = 46.85, p < .001$, with congruent trials, on average, being responded to 15 ms faster than incongruent trials. Furthermore, a significant main effect of target class emerged, $F(1, 13) = 4.89, p < .05$, indicating that trials with object targets were responded to 15 ms slower than trials with animal targets. Finally, a significant interaction effect between target class and congruency became apparent, $F(1, 13) = 11.59, p < .01$. This interaction indicated that the congruency effect was more apparent in trials where the target was an object compared to animal targets (21 ms vs. 6 ms). Posthoc tests revealed that although the congruency effect was larger for the object targets, the congruency effect for both object targets and animal targets was signifi-

Table 2. Median RTs (*SD*) and mean percent error rates (*SD*) as a function of congruency and target class for Experiment 2, and the differences (RCEs) between congruent and incongruent trials (incongruent – congruent)

	Median RTs (<i>SD</i>)			Error rates (<i>SD</i>)		
	Congruent	Incongruent	Difference	Congruent	Incongruent	Difference
Experiment 2	560.0 (45.8)	575.0 (45.9)	15.0	4.3 (2.5)	5.1 (2.8)	0.8
Target class:						
– animal targets	555.1 (51.8)	561.6 (51.4)	6.5	5.6 (3.3)	6.1 (3.8)	0.5
– object targets	562.6 (43.1)	583.4 (43.6)	20.8	3.0 (2.8)	4.2 (2.4)	1.2

Note. Experiment 2: classification of word targets as animals or objects.

cant, respectively, $t(13) = -6.19$, $p < .001$ and $t(13) = -2.74$, $p < .05$.

The same repeated measures analysis performed on error rates only revealed a significant effect of target class, $F(1, 13) = 10.81$, $p < .01$. Accuracy on trials where the target was an animal was less compared to object target trials (5.8% vs. 3.6%). The effect of congruency was not significant, $F(1, 13) = 3.05$, $p > .05$. Although nonsignificant, a trend was still visible: Incongruent trials were more error prone than congruent trials (5.1% vs. 4.3%).

Prime Visibility

Analyses of the posttest showed that on average the participants could only classify 51% of the primes correctly. Again, a direct measure of prime visibility was obtained by treating one level of the response category (i.e., animals) as signal and the other level (i.e., objects) as noise. Mean d' was 0.073. A t -test against the null mean indicated that the mean observed d' value was not significantly different from 0, $t(13) = 0.96$, $p > .35$, indicating that the parameters were sufficient to guarantee unconscious presentation of primes. Using linear regression, a nonsignificant correlation was found between the d' measure and an index for the amount of response priming, $r = .05$, $F(1, 12) = 0.04$, $p > .85$.

Discussion

In Experiment 2 we used a large stimulus set and we adapted the task to resemble the experiment of Forster et al. (2003) more closely. In line with Experiment 1, the data of Experiment 2 again revealed a significant RCE. Although this RCE was more pronounced for the object category, the congruency effect for both object targets and animal targets was significant. Even though no congruency effects were found in the error analysis, the general pattern of results provides further support for the hypothesis that response congruency priming is independent of category size, contradicting the category search model. Furthermore, the direct motor specification hypothesis is also not supported: RCEs were obtained even though novel primes were used and the stimulus set was large.

Experiment 3

Method

Participants

Sixteen Educational Sciences students participated in the experiment as partial fulfillment of a course requirement. One subject was omitted from further analyses because he made substantially (+2*SD*) more errors than the other subjects. Thus, the final sample consisted of 15 subjects (5 male, 10 female) with a mean age of 18.60 (range 17–22).

Apparatus, Stimuli, and Procedure

Apparatus, stimuli, and procedure were exactly the same as in Experiment 2. However, the task changed. Participants were now asked to categorize the targets as being an animal (exemplar) or not (nonexemplar).

Results

RT and Error Rate Analysis

Wrong responses (on average, 4.5% of the data) were discarded. Median RTs from correct responses were entered in a repeated measures analysis with congruency (congruent or incongruent) and target class (animal or nonanimal) as within-subject factors. Median RTs and error rates as a function of congruency and target class are listed in Table 3.

We found a significant main effect of congruency, $F(1, 14) = 11.46$, $p < .01$, with congruent trials being responded to 11 ms faster than incongruent trials. Additionally, a significant main effect of target class emerged, $F(1, 14) = 6.23$, $p < .05$, indicating that trials with nonanimal targets were responded to 18 ms slower than trials with animal targets. No significant interaction effect was found.

The same repeated measures analysis performed on error rates revealed a significant effect of congruency, $F(1, 14) = 6.80$, $p < .05$. More errors were made on incongruent trials compared to congruent trials (5.1% vs. 3.7%).

Table 3. Median RTs (*SD*) and mean percent error rates (*SD*) as a function of congruency and target class for Experiment 3, and the differences (RCEs) between congruent and incongruent trials (incongruent – congruent).

	Median RTs (<i>SD</i>)			Error rates (<i>SD</i>)		
	Congruent	Incongruent	Difference	Congruent	Incongruent	Difference
Experiment 3	570.8 (44.4)	582.0(40.3)	11.2	3.7 (2.9)	5.1 (2.8)	1.4
Target class:						
– animal targets	558.5 (40.8)	576.3 (40.5)	17.8	5.1 (5.4)	5.8 (4.5)	0.7
– nonanimal targets	581.5 (54.1)	588.8 (45.2)	7.3	2.3 (2.3)	4.5 (3.9)	2.2

Note. Experiment 3: classification of word targets as animals or nonanimals.

Prime Visibility

Analyses of the posttest showed that, on average, the participants could only classify 49% of the primes correctly. Again, a direct measure of prime visibility was obtained by treating one level of the response category (i.e., animals) as signal and the other level (i.e., nonanimals) as noise. Mean d' was -0.046 . A t -test against the null mean indicated that the mean observed d' value was not significantly different from 0, $t(14) = -0.46$, $p > .65$, indicating that the parameters were sufficient to guarantee unconscious presentation of primes. Using linear regression, a nonsignificant correlation was found between the d' measure and an index for the amount of response priming, $r = -.29$, $F(1, 13) = 1.17$, $p > .25$.

Discussion

In Experiment 3 we used a large stimulus set and we adapted the task instructions to replicate the experiment of Forster et al. (2003). The findings of Experiment 3 are in line with the results from Experiments 1 and 2, and provide support for the semantic categorization hypothesis. As before, a significant RCE emerged for both RT and error analyses. These results again contradict the direct motor specification hypothesis and the category search model. Contrary to Experiment 2, this RCE was more pronounced for the animal category, although this interaction did not reach significance. This indicates that nonexemplars (nonanimals) do not exhibit consistently larger RCEs than exemplars (animals), as would be predicted by the category search model.

General Discussion

The goal of this study was to clarify the influence of category size on RCEs using novel primes. Regarding this issue, three competing theories have been proposed. In a first theory, proposed by Dehaene et al. (1998), it is assumed that subjects unconsciously apply the task instructions to the primes and, thus, categorize the primes at a semantic level. Therefore, this semantic categorization hypothesis

would predict that RCEs would extend to the novel primes and would be obtained irrespective of category size. A second theory, the direct motor specification hypothesis (Neumann & Klotz, 1994), was implemented in this context by Damian (2001). According to this model semantic processing is bypassed since subjects can automatically associate stimuli with the adequate responses when these stimuli are repeatedly presented to them (e.g., when prime and target set consist of the same stimuli) or when the stimulus set is limited. Thus, this model assumes that when prime and target set consist of different stimuli or when a large stimulus set is used no congruency priming effects would be found. Finally, Forster (2004) utilized the category search model (Landauer & Freedman, 1968) to explain the influence of category size on RCEs. This model states that for broad categories it would be highly unlikely to observe congruency priming effects, since an exhaustive search of the entire category is impossible. However, spurious recognition of some exemplar primes (where an exhaustive category search is often not required) might be obtained and can, therefore, result in a congruency effect for nonexemplar targets. Still, nonexemplar primes will never be recognized as such and, hence, will never produce congruency effects. It is clear that the category search model hypothesizes that when broad categories are used no response congruency effects would be found.

The results of the present study are crystal clear: In all three of our experiments reliable RCEs were obtained. In a first experiment, using two small categories (numbers and body parts) and one broad category (animals), response congruency priming was obtained for all three categories for both RTs and errors. Interestingly, the RCEs were most pronounced for the broad category. In the second experiment, using a larger stimulus set and a task that more closely resembled the experiment of Forster et al. (2003), subjects had to decide whether targets were animals (broad category) or objects (broad category). Again, analyzing the RTs, the observed congruency effects for both objects and animals were significant and a trend toward congruency was present in the error pattern. Finally, the experiment of Forster et al. (2003) was replicated: Subjects had to decide whether targets were animals (exemplars) or not (nonexemplars). As before, a significant RCE emerged in both RT and error analyses. This RCE was not more outspoken for the nonexemplar category (nonanimals).

Summarizing these results, we can conclude that neither the direct motor specification hypothesis nor the category search model can explain these findings. We showed that response congruency priming occurred even when novel primes and a large stimulus set were used, which is not in line with the direct motor specification hypothesis and the claims of Damian (2001). Furthermore, we found RCEs irrespective of category size, which is not in line with the category search model and the claims of Forster (2004). Our findings provide further support for the semantic categorization model proposed by Dehaene et al. (1998): Congruency priming can be observed when novel primes are used (see also Naccache & Dehaene, 2001), irrespective of stimulus set size (see also Reynvoet et al., 2005), irrespective of category size (see also Dell'Acqua & Grainger, 1999), and across different forms of tasks and task instructions.

Although our results are clear, a few comments can be made. First, while our results provide clear evidence in favor of a semantic categorization of unconscious stimuli, this doesn't imply that nonsemantic influences are completely nonexistent. For instance, it seems very likely that automatic stimulus-response mappings may come into play when repeated primes and a small stimulus set are used (as in the experiments of Dehaene et al., 1998). Damian (2001) clearly found that, under these circumstances, RCEs "build up" across an experiment: The RCE showed a clear monotonous rise across blocks and, thus, across increasing repetitions. We believe that such nonsemantic stimulus-response mappings may play a fortifying role in the generation of RCEs under certain circumstances.

Second, the question as to the origin of the RCE still remains. Damian (2001) mentioned the potential confounding of RCEs by semantic relatedness. This occurs in our Experiments 2 and 3: When judging whether a stimulus is an animal or an object/nonanimal, words that are semantically more related also evoke the same response. For example: The masked prime "ape" is conceptually more related to the target "cat" and both evoke the same response, while "ape" is less semantically related to "pot" and both evoke different responses. In this case, one cannot be sure whether the RCE stems from the (in)congruency in evoked responses or whether it is merely a result of the semantic relatedness. However, although this confounding between RCE and semantic relatedness potentially could have contributed to the results for Experiment 2 and 3, it could not have for Experiment 1. Here, the features magnitude, size, and location on the body had to be judged. The RCE is now unconfounded by semantic relatedness. For example: The masked prime "fly" is no more conceptually related to the target "cat" than the target "lion," despite the fact that the first pair is response congruent (both evoke the response "smaller than a Labrador") and the second pair is response incongruent (the prime is smaller than a Labrador while the target is larger than a Labrador.). Reynvoet et al. (2005) also reported re-

liable RCEs when controlling for semantic relatedness as a possible confounder. Furthermore, regardless of whether the origin of the RCE lies in the semantic relatedness or in the (in)congruency of responses, it is undoubted that semantic activation takes place.

Third, how can we reconcile our present findings with the results of Forster et al. (2003, Experiment 3), who found no congruency effect for the animal targets (exemplars), although they did find a significant RCE for the nonanimals (nonexemplars)? It is clear that these differences cannot be attributed to methodological discrepancies: First, the task instructions of Forster et al. were replicated in our Experiment 3, where, nevertheless, reliable RCEs were obtained for both exemplars and nonexemplars. Second, in the present study we used an even stricter masking procedure than the one employed by Forster. In a pilot study, we used a prime duration of 50 ms in order to closely replicate Forster et al.'s (2003) design, where a prime duration of 55 ms was used without backward masking. However, objective prime visibility measures indicated that subjects were able to perceive the primes above chance level. Therefore, we lowered the prime duration to 33 ms. Furthermore, we made use of backward masking of the primes whereas Forster et al. did not. On average, the visibility of the primes in our experiments ranged from 49% to 51%, hence, almost perfectly approaching chance level. Using the paradigm of Forster et al., correct prime classification percentages up to 59% have been previously reported for prime awareness tests (see Lachter, Forster, & Ruthruff, 2004), indicating that our paradigm is even more rigorous.

If methodological inconsistencies are not the origin of the differences between our results and the findings of Forster et al. (2003), what is? We can only speculate about the failure of Forster et al. to observe significant congruency effects for the exemplars (animals). Given the present and previous results (Dell'Acqua & Grainger, 1999; Reynvoet et al., 2005) and the fact that they *did* find significant congruency effects for the nonexemplars, it seems possible that they accidentally obtained null effects for the exemplars. However, using these null effects for the exemplars to propose a theory that refutes that congruency effects can be obtained for broad categories seems premature. The present findings confirm that one should be hesitant, indeed, to jump to conclusions based on a few observations that seemingly fail to fit the reigning theories.

Acknowledgments

Eva Van den Bussche and Bert Reynvoet are affiliated with the Laboratory of Experimental Psychology of the Department of Psychology, University of Leuven and with the Subfaculty of Psychology and Educational Sciences, University of Leuven – Campus Kortrijk. We would like to thank Markus Damian and an anonymous reviewer for their helpful comments and suggestions.

References

- Abrams, R.L., & Greenwald, A.G. (2000). Parts outweigh the whole (word) in unconscious analysis of meaning. *Psychological Science, 11*, 118–124.
- Balota, D.A. (1983). Automatic semantic activation and episodic memory encoding. *Journal of Verbal Learning and Verbal Behavior, 22*, 88–104.
- 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). Imaging unconscious semantic priming. *Nature, 395*, 597–600.
- Dell'Acqua, R., & Grainger, J. (1999). Unconscious semantic priming from pictures. *Cognition, 73*, 1–15.
- Draine, S., & Greenwald, A. (1998). Replicable unconscious semantic priming. *Journal of Experimental Psychology: General, 127*, 286–303.
- Duyck, W., Desmet, T., Verbeke, L., & Brysbaert, M. (2004). WordGen: A tool for word selection and nonword generation in Dutch, German, English, and French. *Behavior Research Methods, Instruments, and Computers, 36*, 488–499.
- 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., 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.
- Greenwald, A.G., Draine, S.C., & Abrams, R.L. (1996). 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 and Cognition, 17*, 35–47.
- 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.
- 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.
- Lachter, J., Forster, K.I., & Ruthruff, E. (2004). Forty-five years after Broadbent (1958): Still no identification without attention. *Psychological Review, 111*, 880–913.
- Landauer, T.K., & Freedman, J.L. (1968). Information retrieval from long-term memory: Category size and recognition time. *Journal of Verbal Learning and Verbal Behavior, 7*, 291–295.
- Marcel, A.J. (1983). Conscious and unconscious perception: Experiments on visual masking and word recognition. *Cognitive Psychology, 15*, 197–237.
- Merikle, P.M., Smilek, D., & Eastwood, J.D. (2001). Perception without awareness: Perspectives from cognitive psychology. *Cognition, 79*, 115–134.
- Naccache, L., & Dehaene, S. (2001). Unconscious semantic priming extends to novel unseen stimuli. *Cognition, 80*, 223–237.
- Neumann, O., & Klotz, W. (1994). Motor responses to nonreportable, masked stimuli: Where is the limit of direct parameter specification? In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance XV* (pp. 123–150). Cambridge, MA: MIT Press.
- Reynvoet, B., Caessens, B., & Brysbaert, M. (2002). Automatic stimulus-response associations may be semantically mediated. *Psychonomic Bulletin and 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.

Received February 14, 2006

Revision received June 13, 2006

Accepted June 30, 2006

Eva Van den Bussche

Faculty of Psychology and Educational Sciences
 University of Leuven
 Campus Kortrijk
 E. Sabbelaan 53
 B-8500 Kortrijk
 Belgium
 Tel. +32 56 246074
 Fax +32 56 246052
 E-mail eva.vandenbussche@kuleuven-kortrijk.be

Appendix A

Stimuli Used in Experiments 1b and 1c (and Their English Translations)

Experiment 1b

Targets	Primes
BIL (buttock)	hiel (heel)
KNIE (knee)	enkel (ankle)
TEEN (toe)	been (leg)
VOET (foot)	kuit (calf)
ARM (arm)	hand (hand)
HOOFD (head)	borst (chest)
BUIK (belly)	nek (neck)
NEUS (nose)	mond (mouth)

Experiment 1c

Targets	Primes
VLO (flea)	muis (mouse)
SPIN (spider)	rat (rat)
KAT (cat)	vlieg (fly)
MUS (sparrow)	mier (ant)
BEER (bear)	paard (horse)
EZEL (donkey)	giraf (giraffe)
PONY (pony)	koe (cow)
LEEUEW (lion)	tijger (tiger)

Note. All possible prime-target combinations were used.

Appendix B

Stimuli Used in Experiments 2 and 3 (and Their English Translations)

Target	Congruent prime	Incongruent prime
AAP (ape)	kat (cat)	pot (pot)
KIP (chicken)	vos (fox)	mes (knife)
MUG (mosquito)	uil (owl)	kar (cart)
RAT (rat)	mus (sparrow)	tol (spinning-top)
VLO (flea)	koe (cow)	bed (bed)
BEER (bear)	rups (caterpillar)	pen (pen)
EEND (duck)	hond (dog)	boog (bow)
EZEL (donkey)	haai (shark)	knop (button)
GIER (vulture)	haas (hare)	hoed (hat)
HAAN (rooster)	gans (goose)	fles (bottle)
HERT (deer)	duif (dove)	touw (string)

Target	Congruent prime	Incongruent prime
KALF (calf)	luis (louse)	bord (plate)
MEES (titmouse)	raaf (raven)	lijm (glue)
MUIS (mouse)	pauw (peacock)	pijl (arrow)
PONY (pony)	mier (ant)	bijl (axe)
RUND (bovine)	spin (spider)	stok (stick)
SLAK (snail)	wesp (wasp)	vlot (raft)
VALK (falcon)	zalm (salmon)	tram (tram)
WOLF (wolf)	egel (hedgehog)	kast (cupboard)
AREND (eagle)	forel (trout)	tafel (table)
GIRAF (giraffe)	havik (hawk)	zetel (chair)
KEVER (beetle)	meeuw (seagull)	plank (board)
LEEUEW (lion)	kraai (crow)	brief (letter)
MEREL (blackbird)	paard (horse)	fluit (flute)
SLANG (snake)	zwaan (swan)	koord (rope)
VLIEG (fly)	stier (bull)	lepel (spoon)
BAL (ball)	pot (pot)	vos (fox)
BUS (bus)	kar (cart)	kat (cat)
KOM (bowl)	mes (knife)	uil (owl)
PAN (pan)	tol (spinning-top)	mus (sparrow)
POP (doll)	bed (bed)	koe (cow)
TAS (bag)	pen (pen)	rups (caterpillar)
AUTO (car)	boog (bow)	hond (dog)
BOEK (book)	knop (button)	haai (shark)
BOOT (boat)	hoed (hat)	haas (hare)
DEUR (door)	fles (bottle)	gans (goose)
GLAS (glass)	touw (string)	duif (dove)
KANO (canoe)	bord (plate)	luis (louse)
KLOK (clock)	lijm (glue)	raaf (raven)
LAMP (lamp)	pijl (arrow)	pauw (peacock)
OVEN (oven)	bijl (axe)	mier (ant)
RING (ring)	stok (stick)	wesp (wasp)
ZAAG (saw)	tram (tram)	spin (spider)
TANG (tongs)	vlot (raft)	zalm (salmon)
VORK (fork)	kast (cupboard)	egel (hedgehog)
VIOOL (violin)	tafel (table)	forel (trout)
PIANO (piano)	zetel (chair)	havik (hawk)
STOEL (chair)	plank (board)	meeuw (seagull)
BEKER (cup)	lepel (spoon)	kraai (crow)
HAMER (hammer)	fluit (flute)	paard (horse)
FIETS (bicycle)	koord (rope)	zwaan (swan)
LAARS (boot)	brief (letter)	stier (bull)