On the Replicability of the Affective Priming Effect in the Pronunciation Task

Adriaan Spruyt,1 Dirk Hermans,1 Mario Pandelaere,1 Jan De Houwer,2 and Paul Eelen1

1University of Leuven, Belgium and 2Gent University, Belgium

Abstract. Bargh, Chaiken, Raymond, and Hymes (1996) and Hermans, De Houwer, and Eelen (1994) showed that a valenced target word is pronounced faster after the presentation of an affectively related prime word than after the presentation of an affectively unrelated prime word. This finding is important because it provides crucial evidence for the hypotheses that stimulus evaluation (a) is goal-independent and (b) facilitates the encoding of stimuli that have the same valence. However, recent studies indicate that the affective priming effect is not a reliable finding in the standard pronunciation task. We report the results of a nearly exact replication of Bargh et al.’s (1996) Experiment 2. In line with previous replication studies, we failed to detect the affective priming effect.

Key words: affective priming, affect, emotion, stimulus evaluation, orthographical depth, semantic activation

Ever since Fazio, Sanbonmatsu, Powell, and Kardes (1986) demonstrated that a valenced target stimulus is responded to faster after the presentation of an affectively related prime stimulus than after the presentation of an affectively unrelated prime stimulus, the affective priming paradigm has been the preferred tool to investigate the conditions under which stimulus valence can be processed. In line with the idea that humans are endowed with an evaluative decision mechanism that allows them to automatically evaluate all incoming stimulus information (e.g., Zajonc, 1980), affective priming studies have demonstrated that stimulus evaluation can occur very rapidly (e.g., Fazio et al., 1986), with minimal effort (Hermans, Crombez, & Eelen, 2000), and outside the reach of consciousness (e.g., Draine & Greenwald, 1998).

There is also evidence that stimulus evaluation can occur even if participants do not have the goal of evaluating stimuli in their environment. Whereas in most affective priming studies participants are instructed to judge the affective connotation of the targets, Bargh, Chaiken, Raymond, and Hymes (1996) asked their participants to pronounce each of a series of target words as fast as possible. Despite the fact that stimulus valence is of no importance in the pronunciation task, it was observed that a target word was pronounced faster when it was preceded by a prime word with the same valence compared to when it was preceded by a prime word with an opposite valence. Hermans, De Houwer, and Eelen (1994) also examined the affective priming effect in the pronunciation task. However, in contrast with the procedure that was used by Bargh et al. (1996), they instructed their participants to pronounce the target words as quickly as possible while ignoring the prime words. Despite this procedural difference,1

1 Another difference between the procedure of Bargh et al. (1996) and Hermans et al. (1994) can be identified: Whereas Bargh et al. (1996) manipulated the prime strength (strong primes versus weak primes), no such manipulation was implemented in the Hermans et al. (1994) study. Note, however, that the affective priming effects that were obtained by Bargh et al. (1996) were not affected by the prime strength manipulation ($P < 1$ in all three experiments).
Hermans et al. (1994) also obtained significant affective priming of pronunciation responses. Taken together, these findings suggest that stimulus evaluation is not conditional on participants having an explicit evaluative goal.

However, the theoretical implications of the pronunciation studies of Bargh et al. (1996) and Hermans et al. (1994) go beyond their support for the hypothesis that stimulus evaluation can occur in the absence of an explicit evaluative goal. These studies are also important because they provide information about the processes that underlie the affective priming effect. Many authors have argued that the affective priming effect is based on a Stroop-like response competition (e.g., Musch, Klauer, & Mierke, 2002). According to this account of affective priming, the primes automatically activate the response alternative that is mapped onto the valence of the targets. For example, when participants are instructed to evaluate the targets, the inferior performance that is observed on incongruent evaluation trials is assumed to be due to the fact that the response alternative that is activated by an affectively unrelated prime differs from the response that should be selected to respond correctly to the target. On the other hand, no such response conflict is assumed to arise when the prime and the target share the same valence, thus giving rise to superior performance. However, in the pronunciation task, the primes and the targets will always elicit different response tendencies because each target is linked with a unique response (i.e., its pronunciation). Therefore, it is difficult to explain affective priming of pronunciation responses on the basis of a response competition mechanism (but see Rothermund & Wentura, 1998). As an alternative explanation, several authors have suggested that the primes preactivate the memory representations of the targets to an extent that the speed and/or the ease with which participants can process their identity is affected (e.g., De Houwer, Hermans, & Spruyt, 2001; Spruyt, Hermans, De Houwer, & Eelen, 2002).

To summarize, the pronunciation studies of Bargh et al. (1996) and Hermans et al. (1994) are important because they provide evidence for the hypotheses that stimulus evaluation (a) is goal-independent and (b) facilitates the encoding of stimuli that have the same valence. However, several studies suggest that the affective priming effect in not a reliable finding in the pronunciation task. For example, in a study that was designed to examine the time course of the affective priming effect in the pronunciation task, Hermans (1996, Experiment 8) failed to detect the affective priming effect at each of five levels of stimulus onset asynchrony (SOA; but see Hermans, De Houwer, & Eelen, 2001). Likewise, De Houwer, Hermans, and Eelen (1998) failed to observe affective priming of pronunciation responses when using non-words of which the meaning had been learned in a preceding learning phase as primes. Significant affective priming was observed only when participants were instructed to categorize the targets as positive or negative.

The most striking null findings were recently published by Klauer and Musch (2001). In a series of pronunciation studies, they repeatedly failed to obtain the affective priming effect. Regardless of the stimulus set size or the SOA that was used, no affective priming effects emerged. Even a nearly exact replication of Bargh et al. (1996) did not yield any significant results. However, because this replication study was carried out in German, it might be argued that linguistic differences between the German and the English language are responsible for the conflicting findings of Bargh et al. (1996) and Klauer and Musch (2001). Indeed, the German and English languages differ in orthographical depth, that is, in the extent to which there is a match between how a word is written and how a word is pronounced (e.g., Frost, Katz, & Bentin, 1987). The German language is characterized by a so-called shallow orthography. In shallow orthographies, the phonemes of a spoken word are represented by the graphemes in a direct and unequivocal manner so that phonology can be determined quickly and efficiently on the basis of orthographic information only. In contrast, the English language is characterized by a deep orthography. In deep orthographies, the orthography-phonology correspondence is more opaque (Frost et al., 1987): The same letter may represent different phonemes in different contexts and different letters may represent the same phoneme (compare, for example, THROUGH and THOUGH). Differences in orthographical depth might be important for affective priming research because it has been repeatedly demonstrated that semantic information exerts the largest impact on the translation from orthography to phonology when this translation is slow or noisy (e.g., De Houwer et al., 2001; Neely, 1991). For example, De Houwer et al. (2001) showed that reliable affective priming of pronunciation responses can be obtained, even in a language with a shallow orthography (i.e., Dutch), when orthographic information is impoverished by placing a percent sign before and after each letter of each target word. Likewise, it has been argued that semantic information exerts a greater influence on the orthography-to-phonology translation in deep orthographies as compared to shallow orthographies because the inconsistent spelling-to-speech correspondences that characterize deep orthographies make the orthography-to-phonology translation proceed more slowly and/or more noisy (Frost et al., 1987; but see Frost & Katz, 1992).

In line with this reasoning, sizeable affective priming effects were readily obtained in pronunciation studies.
that were carried out in a language with deep a orthography (i.e., English; see Glaser & Banaji, 1999; Bargh et al. 1996), whereas rather disappointing results were obtained in pronunciation studies that were carried out in a language with a shallow orthography (i.e., Dutch and German; see De Houwer & Hermans, 1999; De Houwer et al., 1998; Hermans, 1996, Experiment 8; Klauer & Musch, 2001; Spruyt et al., 2002, Experiment 3).

Interestingly, Klauer and Musch (2001) also tested German/English bilinguals with English as well as German stimulus materials. Given the fact that the English language is characterized by a deep orthography whereas the German language is characterized by a shallow orthography, the above reasoning implies that affective priming of pronunciation responses should be more pronounced when German/English bilinguals are tested in English than when they are tested in German. Yet, Klauer and Musch (2001) failed to detect the affective priming effect in both the English as well as the German language. However, as the procedure of their Experiment 4 diverged considerably from the procedure that was used by Bargh et al. (1996), it could be argued that the obtained null finding was due to procedural differences with the procedure that was used by Bargh et al. (1996). In addition, several participants in Experiment 4 of Klauer and Musch (2001) did not acquire the English language as their first language. Therefore, questions can be raised as to what extent the manipulation of the language was sufficiently pure.

For these reasons, we decided to set up a replication study of Bargh et al.'s (1996) Experiment 2 with native speakers of English. The procedure that was used in this experiment was virtually identical to the procedure of Bargh et al. (1996). Nevertheless, two differences with the procedure that was used by Bargh et al. (1996) can be identified. First, we added four meaningless letter strings (e.g., BBB) to the prime set that was used by Bargh et al. (1996, Experiment 2) so that the speed of pronunciation on affectively related and affectively unrelated trials could be assessed relative to a neutral baseline. Note that similar neutral primes were used in Experiment 1 of Bargh et al. (1996). Second, akin to the pronunciation study of Hermans et al. (1994), we instructed our participants to pronounce the targets as quickly as possible while ignoring the primes. In the Bargh et al. (1996) studies, participants were merely instructed to pronounce the second of each pair words.

Method

Participants

Forty-four American students (18 men, 26 women) that studied at various departments of the University of Leuven volunteered to participate. All participants were native speakers of American English.

Materials

Bargh et al. (1996, Experiment 2) used 16 nouns (4 strong-positive, 4 strong-negative, 4 weak-positive, 4 weak-negative) and 20 adjectives (10 positive, 10 negative) as primes and targets, respectively. We used the same set of stimuli for our experiment. Four meaningless three-letter strings (e.g., BBB) were also included in the stimulus set to serve as primes for the control trials.

Bargh et al. (1996, Experiment 2) selected their primes from a larger pool of 92 attitude objects on the basis of the normative study of Bargh, Chaiken, Govender, and Pratto (1992) so that words that were selected to serve as strong primes were associated with the least ambivalent feelings and were evaluated the most extremely, the most quickly, and the most consistently. Likewise, they selected words that were associated with the most ambivalent feelings and that were evaluated the least extremely, the most slowly, and the least consistently to serve as weak primes. In order to assess the reliability of this stimulus selection procedure, we asked our participants, after completion of the affective priming task, to rate the affective extremity of all stimuli on an 11-point rating scale ranging from −5 (very negative) to 0 (neutral) to +5 (very positive). In accordance with the normative data of Bargh et al. (1992), the results of this rating study show that the strong-positive primes, $M = 3.27$ ($SD = 0.49$), were evaluated more positively than the weak-positive primes, $M = 0.98$ ($SD = 0.66$), $t(6) = 4.84, p < .005$, and that the strong-negative primes, $M = -3.83$ ($SD = 0.34$), were evaluated more negatively than the weak-negative primes, $M = -1.22$ ($SD = 0.67$), $t(6) = 6.96, p < .005$. Furthermore, the weak-positive primes were also evaluated more positively than the weak-negative primes, $t(6) = 4.16, p < .01$. Finally, significant differences on the affective dimension were observed between the strong-positive primes and the strong-negative primes, $t(6) = 23.87, p < .005$, and between the positive targets, $M = 3.26$ ($SD = 0.55$), and the negative targets, $M = -3.44$ ($SD = 0.72$), $t(18) = 23.29, p < .005$.

The experiment was run on a MS-DOS IBM compatible Pentium computer. All stimuli were presented
in white uppercase letters (8 mm high, 5 mm wide) against the black background of a 15-inch SVGA computer monitor. Stimulus presentation was controlled by a Turbo Pascal 5.0 computer program. A voice key that stopped a highly accurate Turbo Pascal timer upon registration of a sound (Bovens & Brysbaert, 1990) was used to measure the response latencies.

**Procedure**

All participants were tested individually in a dimly lit room. Participants were told that they were about to participate in a word recognition experiment and that pairs of words would be sequentially presented on the computer screen. They were instructed to pronounce the second of these words as quickly as possible while ignoring the first word. In addition, the use of the voice key was explained and demonstrated in detail.

For each participant, each of the 20 target words was randomly paired with one prime of each prime category. The 100 trials that were thus created were semi-randomly assigned to 5 blocks of 20 trials so that, for each block, 2 primes of each prime category were paired with positive targets and 2 primes of each prime category were paired with negative targets. Prior to the start of the experimental trials, participants were given 10 practice trials (4 congruent trials, 4 incongruent trials, and 2 control trials). These practice trials were identical for all participants but the order of their presentation was randomly assigned to 5 blocks of 20 trials so that, for each block, 2 primes of each prime category were paired with negative targets and 2 primes of each prime category were paired with positive targets. Prior to the start of the experimental trials, participants were given 10 practice trials (4 congruent trials, 4 incongruent trials, and 2 control trials). These practice trials were identical for all participants but the order of their presentation was randomized for each participant. The stimuli that were used to create the practice trials were different from the stimuli of the experimental trials.

Each trial started with a 500 ms presentation of a fixation cross in the center of the computer screen. Five hundred milliseconds after the offset of the fixation cross, the prime words were presented for 200 ms. The target words followed the offset of the prime words after an inter stimulus interval of 100 ms, resulting in a SOA of 300 ms. The target words were displayed until the participant gave a response or 2000 ms elapsed. By pressing one of three keys on the computer keyboard, the experimenter coded whether the microphone was accurately triggered and whether the participant’s response was correct. The next trial was initiated 4000 ms after the experimenter entered the code.

**Results**

Three analyses were carried out. First, we analyzed the data in the same way as did Bargh et al. (1996). Thus, the results of the present experiment can be legitimately compared with the findings of Bargh et al. (1996). Second, also for the sake of comparability, we analyzed the data conform the method that was applied in other publications of the first author of the present work. Finally, we analyzed the test power of the present experiment. Note that the data from the neutral control condition were excluded from the Bargh et al. (1996) analysis as no such prime condition was included in Experiment 2 of Bargh et al. (1996).

**Bargh et al. (1996) Analysis**

The data from the control trials as well as the data from the trials on which the voice key was not appropriately activated (2.10%) or an incorrect response was given (0.74%) were excluded from the analysis. In addition, response latencies less than 300 ms (5.99%) or greater than 1000 ms (0.06%) were also omitted from the analyses. Following Bargh et al. (1996), the raw response latencies were logarithmically transformed prior to computing mean response latencies for each participant and for each experimental condition. The mean (log-transformed) response latencies were then subjected to a $2 \times 2 \times 2$ (Attitude Strength × Prime Valence × Target Valence) analysis of variance with repeated measures for all variables.

The ANOVA revealed a significant main effect of target valence, $F(1, 43) = 5.24, p < .05, MSE = .0035$: Participants were faster to pronounce positive targets ($M = 447$) as compared to negative targets ($M = 453$). However, the crucial interaction between prime valence and target valence failed to reach significance, $F(1, 43) = 1.95, p > .17, MSE = .0019$. All other effects were also non-significant ($ps > .15$).

**Congruency Analysis**

The data from trials on which the voice key was not appropriately activated (2.25%) or an incorrect response was given (0.66%) were excluded from the analysis. In addition, all response latencies that deviated more than 2.5 standard deviations from a participant’s conditional mean latency (2.18%) were also discarded. Next, for each participant and for each experimental condition, mean response latencies were calculated. Mean (untransformed) response latencies were then subjected to a one-way repeated measures ANOVA (congruent trials vs. control trials vs. incongruent trials).

The effect of affective congruency failed to reach significance, $F(1, 43) = 2.11, p > .12, MSE =$
Affective Priming

247.34. A priori contrasts showed that the difference between congruent ($M = 443$) and incongruent priming conditions ($M = 444$) was not statistically significant, $F < 1$. The differences between the control condition ($M = 449$) and the congruent/incongruent priming conditions were also statistically unreliable ($ps > .08$).

Power Analysis

To ascertain that the observed null result was not due to a lack of statistical power, the test power of the present experiment was determined. Given the effect size observed in Experiment 2 of Bargh et al. (1996), $f^2 = 1.13$, the test power of the present experiment, as calculated by the program BWPower of Bakeman and McArthur (1999), is greater than .995. This means that the likelihood of falsely accepting the null hypothesis, $\beta$, is below .005. Even if it is assumed that the effect size of the affective priming effect in the pronunciation task is only one third of the effect size that was estimated on the basis of the results of Bargh et al. (1996, Experiment 2), the probability of falsely accepting the null hypothesis is still below .05. Consequently, for the present experiment, accepting the null hypothesis is as strong a statement as rejecting it.

Discussion

The present experiment was a replication of Experiment 2 of Bargh et al. (1996). It was observed that positive targets were responded to faster than negative targets, an effect that has been observed in previous affective priming studies (e.g., Hermans et al., 1994). This finding is, however, of minimal theoretical importance because various factors that are irrelevant to affective priming research may have produced it. More importantly, we failed to obtain significant affective priming of pronunciation responses. It might be argued that the fact that we instructed our participants to ignore the primes may have interfered with obtaining the affective priming effect. However, as Hermans et al. (1994) did obtain the affective priming effect in the pronunciation task under such conditions, this interpretation of the present null finding seems to be improbable.

The present null finding is in line with many other unsuccessful replication studies (e.g., Klauer & Musch, 2001), and, given the very large test power of our experiment ($1 - \beta > .95$), seriously questions the reliability of the affective priming effect in the pronunciation task. Nevertheless, we believe that the present experiment does not allow one to radically reject the theses that stimulus evaluation (a) is goal-independent and (b) facilitates the encoding of stimuli that have the same valence. Such a conclusion would be difficult to justify because recent studies indicate that reliable affective priming of pronunciation responses can be obtained under certain conditions. For example, Spruyt et al. (2002) recently demonstrated that reliable affective priming of pronunciation responses can be obtained when pictures are used as primes but not when words are used as primes (Experiment 3). This finding is easily explained on the basis of the model of Glaser and Glaser (1989). According to this model, the semantic processing of pictures is far more effective than the semantic processing of words because pictures have privileged access to the semantic system (see also Glaser, 1992). Therefore, as affective information is stored within the semantic system (e.g., De Houwer & Randell, in press), affective priming effects in the pronunciation task are found to be more pronounced when pictures are used as primes as compared to when words are used as primes.

The extent to which the primes are semantically processed is not the only factor that determines whether or not affective priming of pronunciation responses can be observed. The extent to which the targets are semantically processed is also important. De Houwer and Randell (in press) recently showed that reliable affective priming of pronunciation responses can be obtained when participants are instructed to pronounce only those target words that belong to a specific semantic category (e.g., an occupation). The interpretation of this finding is straightforward: Because pronunciation was made conditional upon the detection a specific semantic feature of the targets, the extent to which the targets were semantically processed was substantially increased. As a result, pronunciation responses were considerably affected by the evaluative information that is stored within the semantic system. This interpretation is also in line with the repeatedly demonstrated finding that the magnitude of semantic (and affec-
tive) priming effects increases when target information is degraded (e.g., De Houwer et al., 2001; Neely, 1991). When orthographic information is impoverished, it is beneficial to process the target words semantically in order to determine how to pronounce them correctly. For that reason, the role of semantic (affective) processes in pronunciation is increased when targets are degraded and, hence, significant affective priming of pronunciation responses can be observed under such conditions.

To summarize, recent studies have shown that significant affective priming of pronunciation responses can be obtained whenever the experimental conditions cause participants to semantically process the primes, the targets, or both. As such, these studies indicate that the present null findings do not allow one to radically reject the hypotheses that stimulus evaluation (a) is goal-independent and (b) facilitates the encoding of stimuli that have the same valence. However, at the same time, these studies also suggest that in-depth semantic (affective) processing is unlikely to occur in the standard word-word pronunciation task. The present (statistically reliable) null finding corroborates this interpretation. Of course, this viewpoint only intensifies the question as to what factors were responsible for the semantic (affective) processing that occurred in the word-word pronunciation studies of Bargh et al. (1996) and Hermans et al. (1994). We reasoned that the extent to which semantic (affective) processes are involved in pronunciation might depend upon the orthographical depth of the language in which a word is read, but the present experiment renders this interpretation implausible. Even though we did not explicitly manipulate the orthographical depth, this conclusion is legitimate because a statistically reliable null finding ($\beta < .05$) was observed when testing native speakers of English (a deep orthographic language) with English stimulus materials.

There is, however, one additional factor that may elucidate the present null findings. Some authors have argued that semantic priming effects on pronunciation tend to be obtained only when response latencies are relatively long (e.g., Williams, 1996). Given that the mean response latency of the present experiment was considerably smaller than the mean response latency that was obtained in the Bargh et al. (1996) studies, it could be argued that the response speed was too small in the present experiment for a priming effect to be obtained. For that reason, we reanalyzed the data with response speed (slow responders vs. fast responders) as an additional factor. However, the affective priming effect proved to be unaffected by the factor response speed ($F < 1$). Thus, uncertainty about the factors that are responsible for the conflicting findings that are obtained with the standard word-word pronunciation task continues to exist.

References


Received May 14, 2003
Revision received July 25, 2003
Accepted July 28, 2003

Adriaan Spruyt

Department of Psychology
University of Leuven
Tiestestraat 102
B-3000 Leuven
Belgium
Tel.: +3216325895
Fax: +3216325924
E-mail: Adriaan.Spruyt@psy.kuleuven.ac.be