COGNITION AND EMOTION
2008, 22 (2), 193–217

Psychology Press
Taylor & Francis Group

Revisiting the affective Simon effect

Katia Duscherer
University of Geneva, Geneva, Switzerland
Daniel Holender
Université Libre de Bruxelles, Brussels, Belgium
Esther Molenaar
University of Amsterdam, Amsterdam, The Netherlands

We replicated the affective Simon effect found by De Houwer and Eelen (1998) in a situation in which participants had to respond by saying *positive* or *negative* depending on the grammatical category of the stimulus words while ignoring their affective connotation. Our results show that the affective Simon effect can be modulated by varying the proportion of experimental stimuli bearing a strongly polarised affective connotation. We propose that affective Simon effect depends at least in part on participants' awareness of the correspondence between the affective connotation of the words and the responses. We also submit that this effect might not be specific to affective processing in that it is a token of a vast category of congruity effects that can be based on any kind of meaning of the stimuli, whether semantic or affective.

In the classic spatial version of the *Simon paradigm* as implemented by Craft and Simon (1970), participants respond by a left or a right key press to the green or red colour of a patch while ignoring its left or right position. Faster responses are given with the response key lateralised on the same side as the colour patch, regardless of the colour itself, an observation dubbed the

Correspondence should be addressed to: Katia Duscherer, Faculté de Psychologie et des Sciences de l'Education, 40, Bd. du Pont d'Arve, CH-1211 Geneva 4; Switzerland.
E-mail: kduscher@ulb.ac.be

This work was conducted while KD was a research assistant at the National Fund for Scientific Research – Belgium, and EM a visiting student supported by an Erasmus program at the Laboratoire de Psychologie Expérimentale of the Université Libre de Bruxelles. Part of this study was financed by the Belgian Fund for Joint Basic Research (FRFC, Convention 2.4513.95).

Thanks are due to Jan De Houwer for providing his material and for his helpful comments on an earlier version of the present paper, and to Annette de Groot for helping with the testing in Amsterdam.

www.psypress.com/cogemotion DOI: 10.1080/02699930701339228

spatial *Simon effect*. De Houwer and Eelen (1998; see also De Houwer, 2003a, 2003b; De Houwer, Crombez, Baeyens, & Hermans, 2001a; De Houwer, Hermans, & Eelen, 1998; Tipples, 2001; Voß, Rothermund, & Wentura, 2003, for replications and extensions) reported an affective variant of this paradigm, in which participants respond by saying *positive* or *negative* to words depending on their grammatical category—noun or adjective—while ignoring their positive or negative affective connotation. The *affective Simon effect* corresponds to the observation that participants are faster to respond *positive* than *negative* to a word with a positive affective connotation, and faster to respond *negative* than *positive* to a word with a negative affective connotation. Although both versions of the Simon paradigm are formally equivalent, we suggest that accurate interpretations of the resulting effects require that the specific type of relationships between the stimuli and the responses be taken into account.

Formal analysis of the Simon paradigm

At the formal level, the Simon paradigm can adequately be described by using the taxonomy of tasks proposed by Kornblum, Hasbroucq, and Osman (1990; see also Kornblum, 1992; Kornblum & Lee, 1995). The taxonomy is based on the notion of dimensional overlap and dimensional relevance. Dimensional overlap is defined as "the degree to which two sets of items have properties or attributes in common, and the degree to which such attributes are similar to one another" (Kornblum, 1992, p. 749). Given that there is dimensional overlap at the set-level, the pairing between the elements of two sets can be corresponding or non-corresponding. The notion of dimensional relevance is explained next by showing how it applies to the Simon paradigm.

To apply the notions of dimensional relevance to the Simon paradigm, one has to distinguish between three sets: (a) the *set of relevant stimulus attributes*, which determines the correct responses; (b) the *set of irrelevant stimulus attributes*, which should ideally be ignored; and (c) the *set of responses*. The Simon paradigm is characterised by the fact that the set of relevant stimulus attributes and the other two sets have no dimensional overlap, whereas the set of irrelevant stimulus attributes and the set of responses have. Thus, in the spatial variant of the Simon paradigm, there is no dimensional overlap between the set of relevant stimulus attributes (red or green colour of the stimulus) and either the set of irrelevant stimulus attributes (left or right position of the colour patch) or the set of responses (left or right manual key press). In contrast, there is dimensional overlap between the set of irrelevant stimulus attributes and the set of responses, as the elements of both sets (i.e., the position of the colour patches and the position of the response keys) are lateralised to the left and to the right.

The same formal relations hold in the affective variant of the Simon paradigm. There is no dimensional overlap between the set of relevant stimulus attributes (the grammatical category of the words: noun or adjective) and the two other sets, whereas the set of irrelevant stimulus attributes (positive or negative affective connotation of the words) and the set of responses (the words *positive* or *negative*) have dimensional overlap, as these responses provide obvious category labels for the affective connotations of the stimulus words.

The mapping between the elements of two sets, which have no dimensional overlap, is indifferent to performance: it should be equally slow or fast to respond by a left or a right key press to a colour such as green or red, or to respond by pronouncing the word *positive* or *negative* depending on whether a word is a noun or an adjective. However, when the sets have dimensional overlap, the presentation of an element of one of these sets will automatically activate the corresponding element in the other set (Kornblum, 1992; Kornblum et al., 1990; Kornblum & Lee, 1995). Thus, the dimensional overlap between the set of irrelevant stimulus attributes and the set of responses in the Simon paradigm implies that a colour patch appearing on the right side of the screen activates a right-sided response or that a word with a negative affective connotation like *ugly* activates the response word *negative*. Hence, in *congruent* trials, both the relevant and the irrelevant stimulus attribute evoke the same response; in *incongruent* trials, the relevant and the irrelevant stimulus attribute evoke different responses.[1]

However, the formal analysis of the Simon paradigm does not translate directly into a processing model. Another requirement for a Simon effect to occur is for there to be temporal overlap between the processing of the relevant and irrelevant attributes (e.g., Holender, 1992; Hommel, 1993; Lu, 1997; Lu & Proctor, 1995, 2001). In addition, the time needed for the irrelevant stimulus attribute to activate a response in the response set also depends on the strength of the relation between this irrelevant attribute and the responses, which itself depends on the nature of the dimensional overlap existing between the two sets.

The nature of dimensional overlap

In the spatial Simon paradigm, one source of dimensional overlap between the set of irrelevant stimulus attributes and the set of responses lies in their similarity in terms of spatial relationship: both the stimuli and the response keys are located either on the left or on the right side. Processing the spatial

[1] We will follow De Houwer and Eelen (1998) in using the terms of congruent and incongruent trials instead of consistent and inconsistent trials as in the terminology of Kornblum and Lee (1995).

location of both the stimuli and the response keys can thus be sufficient to induce a spatial Simon effect, an effect that can then be interpreted as reflecting a tendency to react toward the spatial location of a stimulus. This interpretation seems to be supported by observations that the spatial Simon effect can be obtained with monkeys (Riehle, Kornblum, & Requin, 1997) and pigeons (Urcuioli, Vu, & Proctor, 2005), and that, with human participants, the early component of the lateralised readiness potential is determined by the correspondence between the spatial positions of the stimulus and the response, irrespective of whether this response is defined as correct or incorrect through the instructions (De Jong, Liang, & Lauber, 1994, Experiment 4). A secondary source of dimensional overlap between the set of irrelevant stimulus attributes and the set of responses can emerge when the elements of the set of irrelevant stimulus attributes and the set of responses refer to the same category labels. For example, in the spatial Simon paradigm, the spatial relations between the lateralised stimuli and the responses can be specified verbally through the use of descriptors such as *right* and *left*. Hence, with human participants, the effects stemming from the verbal description of the relative spatial positions of the stimuli and the responses may supplement or even completely override the effects stemming from the spatial locations themselves (i.e., Hedge & Marsh, 1975). By using specific descriptions of the different aspects of the experimental situation, the experimenter can thus establish novel relations even between sets with no pre-existing dimensional overlap.

The nature of dimensional overlap between the irrelevant aspects of the stimuli and the responses is, however, not strictly comparable in each possible variant of the Simon paradigm. In our opinion, it is necessary to distinguish between a form of dimensional overlap based upon *stored* semantic information, that is, on stimulus denotation, and a form of dimensional overlap based upon information that can be *generated* on the basis of the meaning of the stimuli, that is, on stimulus connotation. An example of a denotation-based variant is the semantic Simon paradigm devised by De Houwer (1998, Experiment 1). In this experiment, participants respond by saying *animal* or *occupation* depending on whether a word is presented in Dutch or in English, while ignoring that the word names an animal or an occupation. The results show a semantic Simon effect: responses are faster when they correspond to the semantic category of the presented word (i.e., it is faster to respond *animal* than *occupation* to a word like *dog*). The presentation of the word *dog* can indeed activate the concept it denotes, defining that this is the name of a mammal with four legs, often of a brown colour, which wags its tail when it is happy, that is, an animal. On the other hand, we do not believe that there is stored information, and hence information that can be automatically activated, about whether the concept denoted by the word *dog* has a pleasant or an unpleasant connotation

(cf. Holender & Duscherer, 2002, 2004). This information has to be generated by the participant, which implies that the affective Simon paradigm is an example of a connotation-based variant. As a consequence, the affective Simon effect, as any other effect derived from a connotation of the experimental stimuli, must be based on conscious mental representations. We suggest that in an affective Simon experiment, the affective connotation of the stimuli is recoded into the same verbal descriptions as the responses because participants are aware of the relationship between the irrelevant stimulus attributes and the response words selected by the experimenter. Once the noticing of this correspondence occurs, participants can simply not avoid recoding the affective connotations of the stimuli into the same category labels as those used as responses, that is, into the words *positive* and *negative*. Conversely, if participants do not notice the correspondence between the affective connotations of the stimuli and the responses, there is no reason this affective connotation would activate one of the response words, as there would no longer be dimensional overlap between the two sets. Put differently, we imply that in a connotation-based variant of the Simon paradigm, the only source of dimensional overlap between the set of irrelevant stimulus attributes and the set of responses resides in the common verbal description of their elements. If our hypothesis is correct, we should be able to modulate the affective Simon effect by making it harder or easier for the participants to grasp the critical relations in the experiment.

EXPERIMENTS 1A AND 1B

In Experiment 2 of De Houwer and Eelen (1998), participants were explicitly urged not to think about the affective meaning of the words. As a substantial affective Simon effect was observed on the RTs data, the authors concluded that "if the irrelevant feature has an effect, it must have been processed. Because participants were instructed to ignore the irrelevant feature, it can be inferred that the processing of the feature was involuntary and in that sense automatic" (p. 46). We do not dispute this conclusion but if, as suggested above, the affective Simon effect is solely based on participants' noticing and recoding the irrelevant attribute into the categories named by the responses, the adequacy of the instruction to ignore the affective connotation of the words becomes questionable. Indeed, it may have promoted participants' awareness of the critical relationships, and, as a consequence, have enhanced (or even induced) the resulting Simon effect. In Experiment 1a, we explore this possibility by comparing the effects of instructions containing no indication about the meaning of the words with the effects of instructions to ignore the affective connotation of the words.

Other aims are to replicate the affective Simon effect in a language (i.e., French) different from that of the original study (conducted in Dutch, see also De Houwer et al., 2001a), and to assess this effect over a larger corpus of words with strong affective connotations (120 words compared to 40 in the original study).

Method

Participants. Forty undergraduate students at the Université Libre de Bruxelles participated in order to fulfil course requirements. All participants had normal or corrected-to-normal vision, were native French speakers, and were naive about the aim of the experiment. Most of them were in their late teens or early twenties.

Stimuli. Sixty nouns and sixty adjectives were chosen in the affective rating scale of Hermans and De Houwer (1994), established on 740 Dutch words rated by 145 Flemish participants following their affective connotation on a 1 (*very negative*) to 7 (*very positive*) scale. We translated this material into French and selected 120 words that were non-ambiguous in meaning, affective connotation, and grammatical category. As can be seen in Table 1, half of the nouns and half of the adjectives had a strong positive affective connotation and the other halves a strong negative connotation. All words were 1 to 4 syllables long and contained between 3 to 10 letters.

TABLE 1
Mean affective ratings with the corresponding standard deviations (in parentheses) for the experimental words used in each experiment

	Adjectives		*Nouns*	
	Positive	*Negative*	*Positive*	*Negative*
Experiment 1a				
Affective rating	5.9 (0.3)	2.0 (0.3)	5.9 (0.5)	1.6 (0.3)
n	30	30	30	30
Experiment 1b				
Affective rating	5.9 (1.0)	1.9 (1.1)	5.9 (1.0)	1.7 (0.8)
n	10	10	10	10
Experiments 2a[a] and 2b				
Affective rating	5.9 (0.3)	1.8 (0.3)	6.0 (0.4)	1.7 (0.3)
n	10	10	10	10

Note: The affective ratings are based on a 1 (*very negative*) to 7 (*very positive*) scale in *Affective and Subjective Familiarity Ratings of 740 Dutch Words* by Hermans and De Houwer, 1994, *Psychologica Belgica*, *34*, 115–139. [a]Experiment 2a contained 60 additional adjectives with a neutral affective connotation ($M = 4.1$; $SD = 0.8$) and 60 additional nouns with a neutral affective connotation ($M = 4.1$; $SD = 0.3$).

Thirty-eight additional words—19 nouns and 19 adjectives—with a neutral affective connotation were selected: 30 to be used as practice and 8 as warm-up trials.

Apparatus. The experiment was designed using the Micro Experimental Laboratory (MEL II) software (Schneider, 1988). The stimuli were presented on a 70 Hz Highscreen LE48P display controlled by a 486 IBM compatible computer. Verbal responses latencies were recorded in milliseconds using a voice key connected via a MEL II PST Serial Response Box.

Design and procedure. Participants were told that nouns and adjectives would be presented on the screen and that they had to classify these words according to their grammatical category. They were asked to respond fast and accurately. For the 20 participants in the *neutral instruction* condition, there was no information whatsoever about the affective connotation and the meaning of the words. In contrast, the 20 participants in the *directed instruction* condition were explicitly told that each word had a strong affective connotation and that they should try to ignore the affective value of the words while doing the classification task. Half the participants were instructed to say the word *positive* when they saw a noun, and the word *negative* when they saw an adjective, and vice versa for the other half.

On each trial, a single word was presented at the centre of the screen until the voice key recorded a response or for a maximum of 2000 ms. Then, the stimulus vanished and was followed by a black screen for 2000 ms. Errors, ambiguous responses, or extraneous vocal activity were recorded online for later editing of the data set. Words were presented in light-grey on a black background, using the lowercase standard font of the computer. At a viewing distance of 60 cm, the visual angles subtended by each letter were 0.52 degrees in height and 0.31 degrees in width, the gap between two letters being 0.1 degree.

The 120 experimental words were presented a first time in a fixed (i.e., the same for each participant) pseudorandomised sequence constructed with the constraints that there were never more than three consecutive words of the same affective connotation or of the same grammatical category. This was followed by a second presentation of the same 120 words with a different randomisation. A short pause was introduced after completion of 60 trials in each sequence. Each participant was tested individually in one session of about 40 minutes consisting of one practice block of 30 trials and 4 blocks of 60 experimental trials, each beginning with 2 warm-up trials. After the experiment, through a questionnaire we assessed whether participants were aware of the aim of the experiment and to what degree they paid attention to the affective connotations of the words.

Results

Trials in which a voice key failure occurred were not taken into account. For each participant, trials with RTs superior or inferior to two standard deviations from the mean overall RT were discarded. This procedure entailed the elimination of 4.5% of the trials. The mean RTs and error rates for each condition were computed on the remaining trials; they are shown in Table 2. Table 3 shows the mean Simon effects, computed by subtracting the mean RT or the mean error rate for the congruent trials from those of the incongruent trials.

An analysis of variance (ANOVA) with Instruction (neutral, directed) as a between-participants variable and Congruence (congruent, incongruent) and Presentation (first, second) as within-participants variables was

TABLE 2
Mean reaction times (in ms) and error percentages with the corresponding standard deviations (in parentheses) for each experiment and for Experiment 2 of De Houwer and Eelen (1998)

	Experimental trials			
	Presentation 1		*Presentation 2*	
	Congruent	*Incongruent*	*Congruent*	*Incongruent*
Experiment 1a: Neutral instructions				
RTs	837 (85)	845 (95)	808 (85)	826 (104)
Errors	3.7 (3.4)	3.7 (3.3)	2.1 (2.5)	3.8 (3.3)
Experiment 1a: Directed instructions				
RTs	815 (78)	830 (80)	785 (83)	802 (77)
Errors	3.7 (2.8)	2.9 (2.8)	3.0 (3.0)	2.2 (2.1)
Experiment 1b: Directed instructions				
RTs	837 (115)	853 (91)	766 (93)	788 (86)
Errors	5.6 (7.2)	9.1 (5.4)	0.1 (0.1)	5.2 (5.9)
Experiment 2a: Neutral instructions, small proportion				
RTs	828 (112)	825 (114)	—	—
Errors	3.1 (4.6)	7.2 (6.0)	—	—
Experiment 2b: Directed instructions, large proportion				
RTs	837 (131)	865 (137)	797 (105)	819 (111)
Errors	6.8 (11.4)	8.4 (11.0)	4.3 (5.3)	5.6 (4.7)
Experiment 2 of De Houwer and Eelen (1998)[a]				
RTs	723 (147)	788 (166)	703 (125)	731 (148)
Errors	7.2 (—)	7.5 (—)	4.1 (—)	6.0 (—)

Note: [a]Results of Experiment 2 (with directed instructions) of De Houwer and Eelen (1998). An affective variant of the Simon paradigm. To achieve a homogenous presentation of the data, we transformed the mean number of errors, reported in the original article, into percentages.

TABLE 3
Affective Simon effects as a function of participants awareness of the experimental manipulation, with the corresponding .95 standard errors in parentheses

		RTs			*Errors*		
Participants	*n*	*Present. 1*	*Present. 2*	*Overall*	*Present. 1*	*Present. 2*	*Overall*
Experiment 1a: Neutral instructions							
Aware	14	13 (7.2)	18 (8.6)	15 (7.0)	−0.1 (1.0)	1.4 (0.9)	0.6 (0.8)
Unaware	6	−2 (8.4)	19 (9.6)	8 (3.7)	0.3 (1.5)	2.3 (0.9)	1.3 (0.8)
Overall	20	9 (5.8)	18 (6.5)	13 (5.0)	−1.1 (0.8)	1.7 (0.7)	0.9 (0.6)
Experiment 1a: Directed instructions							
Aware	16	17 (7.2)	17 (7.1)	17 (6.1)	−0.5 (0.9)	−0.8 (0.8)	−0.6 (0.6)
Unaware	4	8 (12.4)	16 (5.2)	12 (6.3)	−1.7 (1.1)	−1.3 (1.3)	−1.5 (0.7)
Overall	20	15 (6.2)	17 (5.7)	16 (5.0)	−0.7 (0.8)	−0.9 (0.8)	−0.8 (0.5)
Experiment 1b: Directed instructions							
Aware	16	22 (13.7)	15 (11.2)	18 (10.9)	2.1 (1.8)	1.4 (2.1)	1.8 (1.4)
Unaware	4	−7 (17.5)	50 (14.1)	21 (13.2)	8.9 (5.9)	−1.2 (2.4)	3.9 (2.4)
Overall	20	16 (11.6)	22 (9.8)	18 (9.0)	3.5 (1.9)	0.9 (1.7)	2.2 (1.2)
Experiment 2a: Neutral instructions, small proportion							
Aware	11	5 (14.0)	—	—	1.1 (2.1)	—	—
Unaware	13	−11 (18.4)	—	—	6.8 (1.6)	—	—
Overall	24	−3 (11.7)	—	—	4.1 (1.4)	—	—
Experiment 2b: Directed instructions, large proportion							
Aware	19	34 (14.9)	37 (13.0)	35 (12.6)	1.2 (2.3)	1.8 (1.9)	1.5 (1.8)
Unaware	5	5 (20.2)	−33 (40.0)	−13 (28.2)	4.8 (4.9)	−0.4 (2.9)	2.2 (2.8)
Overall	24	28 (12.6)	22 (14.1)	25 (12.0)	2.0 (2.1)	1.4 (1.6)	1.6 (1.6)

Note: The affective Simon effects are computed by subtracting the mean RTs and the mean error percentages of the congruent trials from those of the incongruent trials. Present. = presentation.

performed. There was a significant main effect of congruence, $F(1, 38) = 17.56$, $MSE = 498.68$, $p < .001$. The main effect of instructions did not reach significance, $F(1, 38) = 0.64$, $MSE = 28{,}098$, $p = .428$, and did not interact with congruence, $F(1, 38) = 0.27$, $MSE = 498.68$, $p = .607$, the congruity effect being of 13 ms for the neutral instruction condition and of 16 ms for the directed instruction condition. The main effect of presentation was significant, $F(1, 38) = 30.27$, $MSE = 921.40$, $p < .001$, with RTs being slower overall for the first ($M = 832$ ms) than for the second ($M = 805$ ms) presentation. Neither the Presentation × Congruence nor the Presentation × Congruence × Instruction interactions were significant, $F(1, 38) = 1.35$, $MSE = 235.08$, $p = .253$; $F(1, 38) = 0.63$, $MSE = 235.08$, $p = .433$.

The same ANOVA was conducted on the error rates. There was no difference between congruent (3.1%) and incongruent (3.2%) trials, $F(1, 38) = 0.01$, $MSE = 6.51$, $p = .942$. Significantly more errors were made during the first presentation than during the second one, $F(1, 38) = 4.89$,

$MSE = 4.67$, $p < .050$. Although the main effect of instruction did not reach significance, $F(1, 38) = 0.32$, $MSE = 18.24$, $p = .578$, the Congruence × Instruction interaction just reached significance, $F(1, 38) = 4.21$, $MSE = 6.51$, $p = .050$. However, subsequent t-tests showed that the difference in error rate between the congruent and incongruent trials did not reach significance in either instruction condition.

Questioned after the experiment, 14 participants in the neutral and 16 participants in the directed instruction condition reported that they were aware of the existence of congruent and incongruent trials. In reconducting the ANOVAs on these subgroups of participants, we found the same pattern of results as for the full groups both for the mean RTs and the error rates. Similar analyses were performed on the 6 and 4 participants claiming unawareness of the presence of congruent and incongruent trials in each condition. For the mean RTs, only the main effects of presentation and congruence were significant, $F(1, 38) = 9.61$, $MSE = 1549.04$ and $F(1, 38) = 9.39$, $MSE = 110.04$; $p < .050$ in each case. None of the effects reached significance in the analysis of the mean error rates.

Discussion

First, we were able to replicate the affective Simon effect found by De Houwer and Eelen (1998) on the RTs data, even though the effect observed in the present study was smaller[2] than the original effect observed by De Houwer and Eelen in their Experiment 2, especially in Presentation 1 (see Table 2). In order to test whether larger effects could be found with our French material, we selected the 40 words yielding the largest positive Simon effect in Experiment 1a and took an additional group of 20 participants in the directed instruction condition. This experiment, which is referred to as Experiment 1b in Tables 1, 2, and 3, had thus the same number of trials as Experiment 2 of De Houwer and Eelen instead of three-times more in Experiment 1a. As can be seen in Table 3, the 16 and 22 ms effects in the first and second presentations were of the same order of magnitude as those of Experiment 1a,[3]

[2] Also, there is a large variability in the effect. Averaged over words, the extreme values of the Simon effect were of −21 ms and +75 ms for participants. Averaged over participants, the extreme values of the effect ranged from −110 to +114 ms for words. Each word functions as its own control through the response assignment conditions, appearing for half of the participants as a congruent and for the other half as an incongruent trial.

[3] In the RT analysis, there were significant main effects of congruence, $F(1, 19) = 4.52$, $MSE = 1606$, $p < .050$, and of presentation, $F(1, 19) = 69.51$, $MSE = 1338$, $p < .050$, with a nonsignificant interaction between them. None of the effects reached significance in the analysis of the error rates. Four participants did not notice the relation between the affective connotations of the words and the responses.

suggesting that the magnitude of the Simon effect cannot be increased with our material and our population of participants.

Second, our prediction of an enhanced effect under the directed instruction is not supported. The results of Experiment 1a suggest that the experimental manipulation is conspicuous enough, irrespective of the instructions participants receive. We can thus rule out that the affective Simon effect reported by de Houwer and Eelen (1998, Experiment 2) was induced solely by the specific instructions they used. In the post-experimental interview, a majority of participants reported being aware of the relations between the affective connotation of the words and the responses in both the directed and the neutral instruction conditions. Actually, even the participants who claimed not to be aware of this critical relation showed a full-blown effect in the second presentation, but no effect at all in the first presentation. With only 6 and 4 participants per condition, there is not enough power for this interaction to be revealed in the mixed-model ANOVA. However, in view of the very similar pattern of results showed by the "unaware" participants in each condition of Experiment 1a and by the 4 "unaware" participants in Experiment 1b (see Table 3), we carried out a within-participants ANOVA on the resulting group of 14 participants. This time a significant Presentation $\times$ Congruence interaction was found, $F(1, 13) = 7.32$, $MSE = 372.9$, $p < .05$, indicating that the Simon effect was nil in the first ($M = -1$ ms, $SD = 25$ ms) and substantial in the second presentation ($M = 27$ ms, $SD = 26$ ms). A comparable analysis carried out on the mean error rates showed no significant effect.

In view of these results, it seems questionable that these 14 participants were completely unaware of the relation between the affective connotations of the words and the responses, especially because even 8 participants out of the 40 receiving the directed instructions claimed not to be aware of what was duly explained to them before they started the experiment. One possibility is that their negative answer reflects more their conviction of not being affected by the manipulation than their unawareness of it.

EXPERIMENTS 2A AND 2B

The aim of Experiments 2a and 2b was to assess whether the affective Simon effect can be modulated by participants' degree of awareness of the critical relationships between the set of irrelevant stimulus attributes and the set of responses. We compare the affective Simon effects in two experimental conditions, the first being intended to minimise the opportunity of noticing these critical relationships, and the second being intended to maximise it. As the small sizes of the obtained affective Simon effects with the French word material used in Experiments 1a and 1b are hampering possible modulations

of the effect, Experiments 2a and 2b were conducted in Dutch with the original material of De Houwer and Eelen (1998).

Experiment 2b is a nearly exact replication of Experiment 2 of De Houwer and Eelen (1998), using 100% of words with a strong affective connotation as well as the directed instructions. In contrast, in Experiment 2a, the proportion of words with a strong affective connotation was reduced to 25% by adding 75% of filler words with a neutral affective connotation. Moreover, participants received the neutral instructions. Although we do not expect that the reduction in the proportion of critical items and the neutral instruction are sufficient to prevent all participants from ever noticing the relationship between the response words and the affective connotations of the stimuli, it should certainly make it less likely. Consequently, we predict at least a diminished effect in this condition.

Method

Participants. Forty-eight undergraduate students at the Universiteit van Amsterdam participated in order to fulfil course requirements, 24 in Experiment 2a and 24 in Experiment 2b. All participants had normal or corrected-to-normal vision and were native Dutch speakers. Most of them were in their late teens or early twenties.

Stimuli. For the experimental words in both experiments we used the same material as in Experiment 2 of De Houwer and Eelen (1998; see Table 1 for the affective ratings). All words were 1 to 4 syllables long and contained between 3 to 10 letters. For Experiment 2a, 150 additional words with neutral affective connotations were selected. Thirty of these words were presented as practice trials, and the other 120 were presented as fillers during the experiment proper. Experiment 2b contained no neutral words. Ten words with positive and ten words with negative affective connotations were used as practice trials.

Apparatus. The experiments were run on a Macintosh Plus ED computer, which also recorded the verbal responses latencies in milliseconds through a connected voice key. Stimulus presentation was controlled by a Turbo Pascal program.

Design and procedure. The design and procedure were identical to those of Experiment 1a with the following exceptions.

In Experiment 2a, participants received instructions that contained no information about the affective connotation of the presented words. They were told that the experiment was part of a research project about grammatical processing. Every word was presented just once, in a fixed

pseudorandomised sequence constructed with the constraints that there were never more than three consecutive words of the same affective connotation or the same grammatical category. There was a break after each 40 trials. Each participant was tested individually in one session of about 40 minutes consisting of one practice block of 30 trials, and four blocks of 40 experimental trials (each preceded by one warm-up trial). One trial consisted of the following sequence of events: (a) a fixation cross for 500 ms at the centre of the screen; (b) a black screen for 500 ms; (c) a single word presented at the centre of the screen until the voice key recorded a response or for a maximum of 3000 ms; and (d) a black screen. After each trial, the experimenter encoded the response (positive, negative, no response, or voice-key error); the next trial was initiated 1500 ms after the code was entered.

In Experiment 2b, participants were told to ignore the affective connotation of the presented words. The 40 experimental words were presented twice with a single break between the two presentation blocks, and no warm-up trials were used.

Results

The cut-off procedure, which was identical to that of Experiment 1a, entailed an overall elimination rate of 5.0% of the trials in Experiment 2a and of 4.6% in Experiment 2b. Error rates were calculated separately for each condition on the remaining trials and mean RTs were computed only for correct responses (see Table 2).

In Experiment 2a, the mean RTs for the experimental trials (thus excluding the filler trials) showed no effect of congruence, $t(23) = 0.29$, $p = .778$. Significantly more errors were produced in incongruent (7.2%) than in congruent (3.1%) trials, $t(23) = 2.96$, $p < .01$. Questioned after the experiment, 11 participants out of 24 reported that they were aware of the existence of incongruent trials. An ANOVA with Awareness (aware, unaware) as a between-participants variable and Congruence (congruent, incongruent) as a within-participants variable carried out on the mean RTs showed no significant effect. A similar ANOVA on the error rates showed that both the main effect of congruence and its interaction with awareness were significant, Fs(1, 22) = 9.12 and 5.21, $MSE = 19.44$, $p < .050$ in both cases. As can be seen in Table 3, this interaction means that there is almost no difference between the error rates for the congruent and incongruent trials for the "aware" participants (3.5 vs. 4.4%) but a large difference for the "unaware" participants (2.8 vs. 9.6%).

In Experiment 2b, the mean RTs were analysed using an ANOVA with Congruence (congruent, incongruent) and Presentation (first, second) as within-participants variables. The main effects of congruence,

$F(1, 23) = 4.36$, $MSE = 3443$, $p < .050$, and of presentation, $F(1, 23) = 15.53$, $MSE = 2792$, $p < .010$, were significant, overall RTs being slower for the first (851 ms) than for the second (808 ms) presentation. The Congruence × Presentation interaction was not significant, $F(1, 23) = 0.23$, $MSE = 832.82$, $p = .640$. With the same analysis, no significant effects were found on the error rates, which were of 5.4% for congruent and of 7.0% for incongruent trials. Questioned after the experiment, 19 participants out of 24 reported that they were aware of the existence of congruent and incongruent trials. In reconducting the ANOVAs on the 19 "aware" participants, we found the same pattern of significant effects for the mean RTs, and no significant effect for the error rates.

To assess whether an interaction between the experiments occurred, an ANOVA with Congruence (2 levels) as a within-participants variable and Experiment (2 levels) as a between-participants variable was conducted on the mean RTs over the 48 participants. This analysis was based on the single set of 40 experimental words in Experiment 2a and on the first presentation of the 40 experimental words in Experiment 2b. Although there was neither a significant main effect of congruence, $F(1, 46) = 2.02$, $MSE = 1774$, $p = .162$, nor of experiment, $F(1, 46) = 0.48$, $MSE = 28{,}907$, $p = .490$, there was a marginally significant interaction between both variables, $F(1, 46) = 3.28$, $MSE = 1774$, $p = .076$. However, the unilateral t-test corresponding to our prediction of at least a reduction in the effect in Experiment 2a compared to Experiment 2b was significant, $t(46) = 1.88$, $p < .05$, showing that the 28 ms effect of Experiment 2b simply vanished in Experiment 2a, being -3 ms. In the ANOVA conducted on the error rates, only the main effect of congruence was significant, $F(1, 46) = 5.933$, $MSE = 1.707$, $p < .05$, reflecting the fact that over the two experiments, there were 4.8% of errors for the congruent trials and 7.8% of errors for incongruent trials.

Discussion

The overall level of performance in Experiment 2b was about 100 ms slower than in Experiment 2 of De Houwer and Eelen (1998), of which it is an almost exact replication. The magnitude of the overall Simon effect was also smaller in Experiment 2b, being of 25 ms instead of 47 ms in Experiment 2 of De Houwer and Eelen. Table 2 shows that this difference in the magnitude of the effects between the two experiments is due mainly to the very large 65 ms effect found by De Houwer and Eelen in Presentation 1, compared to 28 ms in their Presentation 2. As can be seen in Table 4, large Simon effects were also found in Experiment 1 of De Houwer and Eelen and in Experiment 2 of De Houwer et al. (2001a). Turning to the error rates, the results of Experiment 2b confirm those of Experiment 2 of De Houwer and Eelen and the corresponding condition of Experiment 2 of De Houwer et al.

in showing that the slightly larger error rate for the incongruent than for the congruent trials was not sufficient to reach statistical significance. Such was also the case in restricting the analysis of the error rates to the subgroup of participants of Experiment 2b claiming awareness of the relation between the affective connotations of the words and the responses (see Table 3).

The main conclusion about Experiment 2b is that we were able to replicate the affective Simon effect initially demonstrated by De Houwer and Eelen (1998). However, the 25 ms effect we got in using the same material and the same language is not much larger than the significant 16 and 18 ms effects we found with a different material and a different language in the directed instruction condition of Experiments 1a and 1b. At present, we have no ready explanation for the fact that the overall response latencies were about 100 ms longer in all our experiments compared to Experiment 2 of De Houwer and Eelen and Experiment 2 of De Houwer et al. (2001a). However, this is an unlikely cause of the smaller affective Simon effect found in our experiments compared to theirs, because in their experiments, the longer the response latencies, the larger the Simon effect (see general discussion).

Let us now examine our main prediction that, relative to Experiment 2b, we expected at least a reduction in the Simon effect in Experiment 2a. The reason for this prediction was that in Experiment 2a, the relation between the affective connotations of the words and the responses was made less salient by the addition of 75% of words having a neutral affective connotation, and by the use of neutral instructions not calling attention to the meaning and the affective connotation of the words. Actually, the Simon effect as measured by the difference between the mean RT for the congruent and incongruent trials disappeared completely, but a significant effect now appeared on the corresponding difference between the error rates. Moreover, this unanticipated affective Simon effect on the error rates was almost entirely due to the 13 participants not claiming awareness of the critical relationship, with the 11 participants claiming awareness of this relationship showing virtually no effect (see Table 3).

We think that it is possible to interpret the whole pattern of results by appealing to the degree of control participants exert over their performance as a function of the extent to which they are aware of the critical experimental manipulation. A preliminary point worth noticing is that, like in most Stroop and Stroop-like tasks, the participants of Experiments 2a and 2b were able to achieve the goal of making very few errors, generally less than 5% overall. Participants of Experiment 2b, on being informed about the critical manipulation by the directed instructions, were able to avoid making more errors on the incongruent than on the congruent trials right from the beginning of the experiment. What they could not avoid, however, was their response latencies being affected by the congruency of the relation between

the affective connotations of the words and the responses. Conversely, participants like those of Experiment 2a, who were caught by surprise in progressively discovering that some trials were incongruent, could not always avoid making erroneous responses on these trials. We attribute the absence of effect on the response latencies to the fact that, while sporadically noticing the presence of congruent and incongruent trials, none of these participants figured out how systematic the manipulation was before the experiment was over. However, relative to participants not claiming any awareness of the manipulation, those getting some understanding of it could already exert sufficient control for the accuracy of their performance to be unaffected by the congruency of the trials.

GENERAL DISCUSSION

The main finding of the present study can be summarised by three points. First, we replicated the affective Simon effect found by De Houwer and Eelen (1998, Experiment 2) in using the directed instructions in which participants were explicitly told to ignore the affective meaning of the words. This held true in Experiment 2b, which is a nearly exact replication of their Experiment 2 using their 40 Dutch words, in the corresponding condition of Experiment 1a based on a new set of 120 French words, and in Experiment 1b using the subset of 40 French words showing the largest effect in Experiment 1a. However, irrespective of presentation, the magnitude of the effect was in the range of 15 to 28 ms in our experiments, whereas it was larger, especially in the first presentation of Experiment 2 of De Houwer and Eelen (see Table 2).

Second, in Experiment 1a, the average affective Simon effect was already at its maximum under the neutral instruction condition, in which no allusion was made to the affective connotation of the words. Contrary to our prediction, there was no increase of the effect under the directed instruction condition in which we assumed that participants' attention could have been attracted more easily, or sooner in the course of the experiment, to the relationships between the response words and the affective connotation of the stimulus words.

Third, in Experiment 2a, in which in addition to the neutral instructions the proportion of words with a strong affective connotation was reduced to 25% instead of 100% in the other experiments, the Simon effect simply vanished on the response latencies. In contrast, a significant effect now appeared on the error rates. On closer examination, this effect was confined to the 13 participants out of 24 not claiming awareness of the critical relation between the affective connotation of the words and the responses.

Two caveats are in order before we proceed in the discussion. First, it is almost certain that the post-experimental interview underestimated the frequency with which participants were aware of the congruency manipulation, especially in the cases where the noticing of the critical relation was still unsystematic. However, this is the only procedure available that guarantees that participants are not modifying their spontaneous way of dealing with the situation under the different instruction conditions. Second, there is a risk that in denying awareness of the manipulation, some participants actually meant that they believed themselves to be unaffected by it. This is the way we interpreted the denial of awareness even from 8 of the 40 participants getting the directed instructions in Experiments 1a and 1b. Future work with better designed interviews should reduce this latter risk.

Table 2 shows that in Experiments 1a, 1b, and 2b, in which 100% of the words were strongly valenced in terms of their affective connotations, and 50% of the trials were congruent and 50% incongruent, the affective Simon effect was confined to the response latencies, with almost no effect on the error rates. Table 3 shows that this pattern of results was obtained right from the first presentation for the participants claiming awareness of the critical manipulation whether in the directed (Experiments 1a, 1b, 2b) or in the neutral instruction condition (Experiment 1a). We interpret this to mean that these participants almost immediately noticed the relationship between the affective connotations of the words and the responses, irrespective of whether their attention was further attracted to the critical relations through the instructions. We suggest that the participants not claiming awareness (about 25% in each group) of the presence of congruent and incongruent trials just took a little longer to reach a full understanding of the manipulation. This seems to be confirmed by the pooled results of the 14 "unaware" participants of Experiments 1a and 1b: they showed no Simon effect at all, either on response latencies or on response accuracy in the first presentation, whereas in the second presentation, they showed the pattern of results typical of participants fully aware of the manipulation, with a strong affective Simon effect on response latencies and no effect on error rates.

If, however, the percentage of strongly valenced words is reduced to 25% by the addition of 75% of words with a neutral affective connotation—thereby reducing the proportion of congruent and incongruent trials to 12.5% each—the overall pattern of results is the mirror image of the preceding one: no Simon effect on the response latencies, but a significant effect on the error rates, which is confined to the participants not claiming awareness of the critical relationship. In the discussion of Experiment 2a, we submitted that these results reflect different degrees of incipient awareness of the critical relationship between the affective connotation of the words and the responses. The smallest degree of awareness was reached by participants showing a Simon effect on the error rates but claiming unawareness of the

manipulation. A somewhat higher degree of awareness enabled the other participants to control for their error rates and to claim awareness of the manipulation. Presumably, with more blocks of trials these participants would have kept their error rates under control, while progressively showing an affective Simon effect on the response latencies.

To summarise, the overall results of the present study suggest that the affective Simon effect can be modulated by participants' awareness of the critical manipulation. This is consonant with our hypothesis stated in the introduction that, in the affective Simon paradigm, the only way a link can be established between the irrelevant aspects of the stimuli (i.e., the affective connotation of the words) and the responses is through a common verbal, propositional description. This hypothesis stems from the distinction we make between stored semantic knowledge, like the underlying concepts the words denote, and knowledge that has to be generated by the participants, like the affective connotation of the same concepts.

Several authors will dispute this last point, claiming that stimulus valence *is* stored in long-term memory and that it can be activated and evaluated independently of other cognitive processes. Some evidence supporting the existence of stored stimulus valences stems from affective priming studies in which participants are faster to pronounce a target word like *happy* when it is preceded by a prime word bearing the same stimulus valence, like *flower*, than when it is preceded by a prime word bearing the opposite stimulus valence, like *war* (e.g., Bargh, Chaiken, Raymond, & Hymes, 1996; De Houwer & Randell, 2004; Hermans, De Houwer, & Eelen, 2001). These effects have proven to be very difficult to replicate (De Houwer, Hermans, & Spruyt, 2001b; Klauer & Musch, 2001, 2003; Spruyt, Hermans, Pandelaere, De Houwer, & Eelen, 2004), but let us assume for the moment that the affective valences of known stimuli are stored and can be activated by the mere exposure of the corresponding words. If all positively valenced stimuli share the same stored positive valence trait and if all negatively valenced stimuli share the same stored negative valence trait, the affective Simon effect could, in theory, be caused by the stimulus word pre-activating an underlying trait of the concept underlying the response word. However, following this line of thought, the presentation of a single valenced word should activate through a single trait a large part of our entire semantic memory. Let alone the low ecological value of such a phenomenon, it seems unlikely that such an overall activation process would result in a response conflict between the specific words "positive" and "negative" as observed in the Simon paradigm. Besides, the fact that the affective Simon effect is modulated by the proportion of stimulus words with a strong affective connotation and by participants' awareness of the correspondence between the stimulus and the response sets does not support an interpretation of the effect in terms of an irrepressible activation process. This being said, we need

to emphasise that we can draw no conclusions on the basis of the present results concerning the actual existence of stored valences of known stimuli, the existence of an automatic appraisal process, or the experience of genuine emotional experiences by the participants of an affective Simon task. The only statement we can make about these questions on the basis of our present results is that the affective Simon paradigm does not seem to be the most appropriate tool for their study, as the resulting effects depend on context factors and participants' awareness of the correspondences in the experimental situation.

One way to test the soundness of our analysis is to look at whether it applies only to the affective version of the Simon paradigm or whether it generalises to other variants as well. Table 4 summarises the results of most experiments relevant to this discussion. Except for Experiment 1 of De Houwer and Eelen (1998), all the experiments were based on binary responses, with 50% of congruent and 50% of incongruent trials (which implies that the critical relation is realised in 100% of the trials), and two presentations of the material. All but two of these experiments rested on the directed instructions, in which participants were explicitly told to ignore the affective connotation of the words. The two exceptions are Experiment 4 of De Houwer et al. (2001a) using the neutral instructions and Experiment 1 of De Houwer and Eelen (1998), in which participants were simply told to ignore the meaning of the words, without specifying anything about their affective connotation. Both in these two experiments and in most of the experiments based on the directed instructions, there were significant effects on the response latencies and small, generally nonsignificant, effects on the error rates. However, the magnitudes of the Simon effects varied considerably between experiments and conditions, ranging from 14 to 66 ms in the 13 significant effects shown in Table 4.

To account for the differences in the magnitude of the Simon effects, it is necessary to analyse the relative speed of processing of the relevant and irrelevant attributes of the stimuli (e.g., Holender, 1992; Hommel, 1993; Lu, 1997; Lu & Proctor, 1995, 2001). In the introduction, the Simon paradigm was characterised by the fact that there is dimensional overlap only between the set irrelevant stimulus attributes and the set of responses. In the absence of dimensional overlap between the set of relevant stimulus attributes and the set of responses, the time needed to map an arbitrary response to the relevant stimulus attribute depends on stimulus format (picture vs. written word), on the nature of the relevant attribute (physical property vs. semantic or grammatical category), and on the nature of the responses (verbal vs. manual). In addition to being affected by these variables, the time needed for the irrelevant stimulus attribute to activate a response in the response set also depends on the strength of the relation between this irrelevant attribute and the responses, which itself depends on the nature of the dimensional

TABLE 4
Synopsis of affective and semantic Simon experiments, with the mean RTs and the mean error rates, for congruent and incongruent trials, and the resulting Simon effects

					RTs (ms)		*Errors (%)*		*Simon effect*	
Study	*Exp.*[a]	*Relevant attribute*	*Irrelevant attribute*	*Responses*[b]	*C*	*I*	*C*	*I*	*RTs (ms)*	*Errors (%)*
De Houwer (1998)	1	Dutch or English	Names of occupations and animals	*occupation, animal*	665	710	2.8	5.3	45*	2.5*
	2	Upper or lower case	Names of occupations and animals	*occupation, animal*	584	605	1.8	1.8	21*	0.0
	3	Noun or adjective	Names of persons and animals	*person, animal*	678	730	1.5	2.1	52*	0.6
De Houwer [c] & Eelen (1998)	1	Noun or adjective	Neutral and affectively valenced words	*positive, negative*	759	825	1.9	4.7	66*	2.9
	2	Noun or adjective	Affectively valenced words	*positive, negative*	713	760	5.6	6.7	47*	1.1
	3	Noun or adjective	Affectively valenced words	*flower, cancer*	768	793	5.0	8.4	25*	3.4†
De Houwer [c] et al. (2001)	1	Person or animal	Affectively valenced words	*positive, negative*	632	663	1.4	2.4	31*	1.0
	2	Upper or lower case; noun or adjective	Affectively valenced words	*positive, negative*	509	524	1.4	1.4	15*	0.0
			Affectively valenced words	*positive, negative*	735	789	3.9	6.4	54*	2.5
	3	Man-made or natural; black and white or colour	Affectively valenced pictures	*positive, negative*	820	855	2.8	2.0	35*	0.8
			Affectively valenced pictures	*positive, negative*	665	667	0.9	1.8	−2	0.9
	4	Noun or adjective	Affectively valenced words	Approach or remove a manikin	987	1020	6.2	7.9	33*	1.7
Tipples (2001)[d]		Human or animal	Affectively valenced pictures	*nice, nasty*	656	694	NR	NR	38*	NR
		Human or animal	Affectively valenced pictures	*comedy, cancer*	762	778	NR	NR	14*	NR

Note: [a]All experiments with the exception of Experiment 1 of De Houwer and Eelen (1998) and Experiment 4 of De Houwer et al. (2001) used the directed instructions. [b]All the experiments used binary verbal responses consisting of the words indicated in the cells, with the exception of Experiment 4 by De Houwer et al. (2001), in which manual responses were used. [c]Mean percentages of errors computed from the reported mean number of errors. [d]Mean RTs and effect sizes estimated from figures. Exp. = Experiment; C = congruent; I = incongruent; NR = non reported. *Significant, $p < .05$; †marginally significant.

overlap existing between the two sets. Four different forms of dimensional overlap between the irrelevant aspects of the stimuli and the responses can be distinguished in Table 4: one form lies in the *denotation* of the stimuli in terms of the semantic category to which they belong; the three other forms lie in the *connotation* of the stimuli in terms of their affective valences.

In the Irrelevant-Denotation case, a large semantic Simon effect of 45 ms was found in Experiment 1 of De Houwer (1998), in which animal names and occupation names were used as stimuli and the responses consisted in pronouncing the category names *animal* or *occupation* according to the language of the words. This effect was reduced to 21 ms in Experiment 2, in which the relevant attribute was the case in which the words were written, and in which the overall speed of responding was nearly 100 ms faster than in Experiment 1.

In the Irrelevant-Connotation 1 case, the words used as responses provide descriptive labels for the valences of the affective meaning of the stimuli. Thus, the words *positive* and *negative* (De Houwer & Eelen, 1998; De Houwer et al., 2001a) and the words *nice* and *nasty* (Tipples, 2001) constitute adequate labels for the strong positive and negative affective valences of the word and the picture stimuli. In these experiments, the largest affective Simon effects were obtained with a grammatical categorisation task, as in Experiments 1 and 2 of De Houwer and Eelen or Experiment 2 of De Houwer et al. In the latter experiment, faster overall responses, but also a reduced Simon effect, were observed when participants responded instead to the letter case of the target words. Similarly, in Experiment 3 of De Houwer et al., the affective Simon effect of 35 ms with the semantic categorisation task completely vanished when the same pictures were responded to according to their colour, with the overall speed of responding being about 170 ms faster in the latter than in the former condition.

Thus, irrespective of whether the irrelevant attribute was the semantic category (Irrelevant-Denotation case) or the affective connotation (Irrelevant-Connotation 1 case) of the stimuli, the difference in the magnitude of the Simon effects in each of the comparisons between tasks just examined can be interpreted entirely in terms of the relative speeds of processing of the relevant and irrelevant stimulus attributes, because the stimuli, the responses, and the nature—and thereby the strength—of the relation between the irrelevant attribute and the responses were kept constant in each of these comparisons. Moreover, even across different word materials and different responses, the slower the task, the larger the Simon effect. In Table 4, this trend can be clearly seen in the tasks consisting in judging the semantic and the grammatical category of the words (De Houwer, 1998; De Houwer et al., 2001a; De Houwer & Eelen, 1998). Therefore, we submit that the major determinant of the magnitude of the Simon effect in the experiments of Table 4 implementing the Irrelevant-Denotation and the

Irrelevant-Connotation 1 cases is the relative speed of processing of the relevant and irrelevant attributes, with the strength of the relation between the irrelevant attribute and the responses being very high, and thereby the speed of processing of the irrelevant attribute being very fast. Note in particular that the link between the affective valences of the words and the responses *positive* and *negative* is strong enough to elicit a small but significant affective Simon effect even in the task of judging the letter case of the words (De Houwer et al., 2001a, Experiment 2), which by far led to the shortest response latencies.

Let us now specify the characteristics of the two other irrelevant connotation cases and show how these cases relate to the Irrelevant-Connotation 1 case. We restrict our comparisons to conditions based on the same stimulus material, but here the responses or even the response modes (verbal vs. manual) are different between conditions.

In the Irrelevant-Connotation 2 case, the relationship is established between the affective connotations of the stimuli and the affective connotations of the responses. Thus, in Experiment 3 of De Houwer and Eelen (1998), which was based on nouns and adjectives with strong positive and negative affective connotations, the responses *flower* and *cancer* also had a strong positive and a strong negative affective connotation, respectively. Relative to the experiments that implemented the comparable Connotation 1 condition by using the responses *positive* and *negative* (Experiment 2 of the same study and Experiment 2 of De Houwer et al., 2001a), the overall speed of responding to the grammatical category of the words was only a bit slower in Experiment 3, but the affective Simon effect was reduced. Another example is the experiment by Tipples (2001), which used affectively valenced picture stimuli and the responses *comedy* and *cancer*. Relative to the condition in which the responses *nice* and *nasty* were used, categorisation responses were nearly 100 ms slower with the responses *comedy* and *cancer*, and the affective Simon effect dropped from 38 to 14 ms.

In the Irrelevant-Connotation 3 case, the relationship established between the affective connotations of the stimuli and the responses is metaphorical. The only example is Experiment 4 of De Houwer et al. (2001a), in which the stimuli were adjectives and nouns with strong positive and negative affective connotations and the responses consisted in moving a manikin towards or away from these words as a function of their grammatical category. Here, the relation between the affective connotation of the words and the responses rests on the metaphorical interpretation of the direction of the responses in terms of approach and avoidance. While this experiment yields the slowest overall speed of responding, there is still a 33 ms affective Simon effect.

We thus suggest that the major determinant of the smaller magnitude of the Simon effect in the Irrelevant-Connotation 2 and the Irrelevant-Connotation 3 cases relative to the Irrelevant-Connotation 1 case lies in

the difference in the strength of the relation between the irrelevant attribute and the responses. The reason is that, in each of the comparisons made above, the slower tasks also yielded the smaller Simon effects. Of course, it is probably an oversimplification to assume that the difference in relative strength of the critical relation only affects the amplitude of activation of the irrelevant stimulus–response association. It can also affect the frequency with which this relation affects performance and the amount of practice needed for this relation to become conspicuous enough to affect performance. Except for the experiment of Tipples (2001), in which the error rates were not reported, all of these experiments display the pattern of results typical of participants who have a full understanding of the critical manipulation, that is large effects on response latencies and negligible effects on the error rates. Moreover, this pattern of results was obtained right from the first presentation in Experiment 3 of De Houwer and Eelen (1998), which implemented an Irrelevant-Connotation 2 case. Hence, it does not seem that the relative strength of the relation between the irrelevant attribute and the responses affects the saliency of this relation, probably because the stimuli always have strong positive and negative affective connotations in all the experiments reported so far.

In conclusion, the analyses made above indicate that the Irrelevant-Denotation and the Irrelevant-Connotation 1 cases share the same properties. If this is correct, there is no ground for distinguishing between a semantic form of the Simon effect based on the irrelevant denotation of the stimuli in terms of their semantic category (De Houwer, 1998) and an affective form of the Simon effect based on the connotation of the stimuli in terms of their affective valence (De Houwer & Eelen, 1998). All the results of Table 4 can be accounted for by general principles that apply to any kind of Simon task (e.g., Hommel, 1993; Lu, 1997; Lu & Proctor, 1995, 2001): the magnitude of the Simon effect depends on the relative speed of processing of the relevant and irrelevant information and on the relative strength of the relation between the irrelevant attribute and the responses.

Naturally there are many connotations, in addition to the affective connotation, that can be assigned to any referent and used as bases for classification (e.g., large or small objects, useful or useless objects, gentle or ferocious animals, natural or artificial landscapes). Like affective valence, these connotations can be derived, that is they can be generated from the knowledge of the meaning of the referents, and they can be used to design a multitude of new variants of the Simon paradigm. However, again, these various connotations do not correspond to stored information that can be accessed by the mere presence of the appropriate stimuli, and any connotation-based variant of the Simon effect cannot be but dependent on participants' conscious understanding and recoding of the relation between the connotations of the stimuli and the responses. So, although

there is a multitude of ways in which the formal relations involved in the Simon paradigm can be implemented, as long as the effect is caused by linguistic mediation, it is doubtful that we can learn anything specific about the processing of the irrelevant attribute.

Manuscript received 29 July 2002
Revised manuscript received 1 March 2007
Manuscript accepted 14 March 2007
First published online 28 August 2007

REFERENCES

Bargh, J. A., Chaiken, S., Raymond, P., & Hymes, C. (1996). The automatic evaluation effect: Unconditional automatic attitude activation with a pronunciation task. *Journal of Experimental Social Psychology*, *104*, 104–128.

Craft, J. L., & Simon, R. (1970). Processing symbolic information from a visual display: Interference from an irrelevant directional cue. *Journal of Experimental Psychology*, *83*, 415–420.

De Houwer, J. (1998). The semantic Simon effect. *Quarterly Journal of Experimental Psychology*, *51A*, 683–688.

De Houwer, J. (2003a). The extrinsic affective Simon task. *Experimental Psychology*, *50*, 77–85.

De Houwer, J. (2003b). A structural analysis of indirect measures of attitudes. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation* (pp. 219–243). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

De Houwer, J., Crombez, G., Baeyens, F., & Hermans, D. (2001a). On the generality of the affective Simon effect. *Cognition and Emotion*, *15*, 289–206.

De Houwer, J., & Eelen, P. (1998). An affective variant of the Simon paradigm. *Cognition and Emotion*, *12*, 45–61.

De Houwer, J., Hermans, D., & Eelen, P. (1998). Affective Simon effects using facial expressions as affective stimuli. *Zeitschrift für Experimentelle Psychologie*, *45*, 88–98.

De Houwer, J., Hermans, D., & Spruyt, A. (2001b). Affective priming of pronunciation responses: Effects of target degradation. *Journal of Experimental Social Psychology*, *37*, 87–91.

De Houwer, J., & Randell, T. (2004). Robust affective priming effects in a conditional pronunciation task: Evidence for the semantic representation of evaluative information. *Cognition and Emotion*, *18*, 251–264.

De Jong, R., Liang, C.-C., & Lauber, E. (1994). Conditional and unconditional automaticity: A dual-process model of effects of spatial stimulus-response correspondence. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 731–750.

Hedge, A., & Marsh, N. W. A. (1975). The effect of irrelevant spatial correspondences on two-choice response-time. *Acta Psychologica*, *39*, 427–439.

Hermans, D., & De Houwer, J. (1994). Affective and subjective familiarity ratings of 740 Dutch words. *Psychologica Belgica*, *34*, 115–139.

Hermans, D., De Houwer, J., & Eelen, P. (2001). A time course analysis of the affective priming effect. *Cognition and Emotion*, *15*, 143–165.

Holender, D. (1992). Expectancy effects, congruity effects, and the interpretation of response latency measurement. In J. Alegria, D. Holender, J. Junca de Morais, & M. Radeau (Eds.), *Analytic approaches to human cognition* (pp. 351–375). Amsterdam: North Holland.

Holender, D., & Duscherer, K. (2002). Unconscious semantic access: A case against the hyperpowerful unconscious. *Behavioral & Brain Sciences*, *25*, 340–341.

Holender, D., & Duscherer, K. (2004). Unconscious perception: The need for a paradigm shift. *Perception & Psychophysics*, *66*, 872–881.

Hommel, B. (1993). The relationship between stimulus processing and response selection in the Simon task: Evidence for temporal overlap. *Psychological Research*, *55*, 280–290.

Klauer, K. C., & Musch, J. (2001). Does sunshine prime loyal? Affective priming in the naming task. *Quarterly Journal of Experimental Psychology*, *54*, 727–751.

Klauer, K. C., & Musch, J. (2003). Affective priming: Findings and theories. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation* (pp. 7–49). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Kornblum, S. (1992). Dimensional overlap and dimensional relevance in stimulus–response and stimulus–stimulus compatibility. In G. E. Stelmach & J. Requin (Eds.), *Tutorials in motor behavior II* (pp. 743–777). Amsterdam: North Holland.

Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: Cognitive basis for stimulus-response compatibility: A model and taxonomy. *Psychological Review*, *97*, 253–270.

Kornblum, S., & Lee, J. W. (1995). Stimulus–response compatibility with relevant and irrelevant stimulus attributes that do and do not overlap with the response: A model and taxonomy. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 855–875.

Lu, C.-H. (1997). Correspondence effects for irrelevant information in choice-reaction tasks: Characterizing the stimulus–response relations and the processing dynamics. In B. Hommel & W. Prinz (Eds.), *Theoretical issues in stimulus–response compatibility* (pp. 85–117). Amsterdam: North Holland.

Lu, C.-H., & Proctor, R. W. (1995). The influence of irrelevant location information on performance: A review of the Simon and spatial Stroop effects. *Psychonomic Bulletin & Review*, *2*, 174–207.

Lu, C.-H., & Proctor, R. W. (2001). Influence of irrelevant information on human performance: Effects of S–R association strength and relative timing. *The Quarterly Journal of Experimental Psychology*, *54A*, 95–136.

Riehle, A., Kornblum, S., & Requin, J. (1997). Neuronal correlates of sensorimotor association in stimulus–response compatibility. *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 1708–1726.

Schneider, W. (1988). Micro-experimental laboratory: An integrated system for IBM PC compatibles. *Behavior Research Methods, Instruments, and Computers*, *20*, 206–217.

Spruyt, A., Hermans, D., Pandelaere, M., De Houwer, J., & Eelen, P. (2004). On the replicability of the affective priming effect in the pronunciation task. *Experimental Psychology*, *51*, 109–115.

Tipples, J. (2001). A conceptual replication and extension of the affective Simon effect. *Cognition and Emotion*, *15*, 705–710.

Urcuioli, P. J., Vu, K. P. L., & Proctor, R. W. (2005). A Simon effect in pigeons. *Journal of Experimental Psychology: General*, *134*, 93–107.

Voß, A., Rothermund, K., & Wentura, D. (2003). Estimating the valence of single stimuli: A new variant of the affective Simon task. *Experimental Psychology*, *50*, 86–96.