Failures to induce implicit evaluations by means of approach–avoid training

Katrien Vandenbosch & Jan De Houwer

Failures to induce implicit evaluations by means of approach–avoid training

Katrien Vandenbosch and Jan De Houwer

Department of Experimental-Clinical and Health Psychology, Ghent University, Ghent, Belgium

Woud, Becker, and Rinck (2008) asked participants to repeatedly push pictures of certain faces away and to pull pictures of other faces towards them using a joystick. Performance in a subsequent affective priming task showed that previously pulled faces evoked more positive implicit evaluations then previously pushed faces. We report five studies in which we failed to find consistent evidence for the effect of approach–avoid training on implicit evaluations. We also failed to reproduce the effect reported by Woud et al. when reanalysing their data. An overall analysis that included our data, the data of Woud et al. and additional data provided by Woud and colleagues also did not reveal a significant effect of approach–avoid training on implicit evaluations. We conclude that the impact of approach–avoid training on implicit evaluations is subject to subtle boundary conditions.

Keywords: Implicit evaluation; Learning; Attitudes.

Ever since the seminal publication of Zajonc (1980), researchers have directed their attention to spontaneous preferences, that is, evaluative responses that are evoked automatically in the sense of unintentional, uncontrolled, unconscious, efficient, or fast (see De Houwer & Hermans, 2010; Ferguson & Zayas, 2009, for reviews). These spontaneous preferences, which we refer to as implicit evaluations (see De Houwer, 2009), have been shown to contribute in a unique manner to many psychological phenomena, including psychopathology (see Roefs et al., 2011), addiction (Wiers & Stacy, 2006), and social interactions (Fazio & Olson, 2003).[1]

Until recently, most research on implicit evaluations aimed at constructing tasks to capture these implicit evaluations, demonstrating the manner in which evaluation can be implicit, or exploring the relation between implicit evaluation and behaviour (see De Houwer & Hermans, 2010, for a review). In a relatively small number of studies, researchers have also started to explore

---

[1] Note that in line with the functional-cognitive approach to psychology (De Houwer, 2009, 2011), we consider implicit evaluations to be behavioural effects (i.e., the automatic impact of stimuli on evaluative responses) rather than mediating mental constructs (i.e., mental representations that mediate the impact of stimuli on evaluative responses).

ways to induce and change implicit evaluations (e.g., Baccus, Baldwin, & Packer, 2004; Woud, Becker, & Rinck, 2008). Given that we now know that implicit evaluations have a unique impact on many types of behaviour, it is indeed important to design methods that allow us to influence them. We report a series of studies that were designed to evaluate the merits of one of these methods, namely the approach–avoid training method that was described by Woud et al. (2008) in this journal.

Woud et al. (2008) reported a single study in which participants saw pictures of male and female faces with a neutral facial expression. During the training task, the pictures had a slight blue or brown colour. Participants were asked to push or pull a joystick in front of them on the basis of the colour of the pictures. When they pushed the joystick, the picture became smaller, creating the illusion that they were pushing the picture away from their body. When they pulled the joystick, the picture became bigger, thus creating the illusion that they were pulling the picture towards their body. Unbeknown to the participants, pictures showing certain faces always had a blue colour whereas pictures showing other faces always had a brown colour. As such, participants repeatedly (10, 20, or 50 times) pushed away certain faces and repeatedly pulled other faces towards them.

After the training task, participants completed an affective priming task and rated their liking of the faces. The affective priming task was designed to capture the implicit evaluations of the training faces. Each face was briefly presented as a prime that was sometimes followed by a positive target stimulus and sometimes by a negative target stimulus. Participants were asked to evaluate the targets as positive or negative. Given that positive primes are known to automatically facilitate the evaluation of positive relative to negative targets and that negative primes automatically facilitate the evaluation of negative relative to positive targets (e.g., Fazio, Sanbonmatsu, Powell, & Kardes, 1986), the impact of the face primes on target responses can be used as an index of the implicit evaluations of the face primes (see Fazio, Jackson, Dunton, & Williams, 1995). Immediately before or after the priming task, participants were also asked to rate all trained faces with regard to how sympathetic they found each face to be. Woud et al. (2008) reported that the approach–avoid training procedure had a strong effect ($\eta^2 = .23$) on the implicit evaluation of the trained faces. Faces that were repeatedly pushed away during training subsequently facilitated responses to negative targets relative to positive targets in the affective priming task whereas the reverse was true for faces that were repeatedly pulled towards the participant during training. Ratings, however, were not affected by the training procedure. Woud et al. thus concluded that their approach–avoid training procedure could be used to induce implicit evaluations.

Given the apparent effectiveness of the training procedure described by Woud et al. (2008), we were eager to replicate their effect in our laboratory and to explore the conditions under which the effect occurs. Unfortunately, in several studies, we failed to replicate the basic effect of approach–avoid training on implicit evaluation. This prompted us to reanalyse the original data of Woud et al., and to analyse various combinations of data sets in order to uncover the boundary conditions of the approach–avoid training effect. In the remainder of this paper, we first describe our attempts to replicate the approach–avoid training effect, followed by the reanalysis of the Woud et al. (2008) data and more recent data from their lab, the overall analysis of all available data and subsets of that data, and a discussion of the implications of our findings. All experiments were based on the procedures used by Woud et al. (see Table 1). Any deviation from these procedures will be mentioned explicitly.

EXPERIMENT 1

Method

Participants. Twenty-nine students (14 women, mean age = 18.52, $SD = 0.85$) at Ghent University took part in the experiment in exchange for course credits. All participants gave their informed consent prior to the start of the experiment.

Table 1. *Procedural differences between experiments*

| | Training task | | | Affective priming task | | | |
Experiment	Attentional manipulation	Joystick movement	Zoom function	Stimulus onset asychrony	Primes involving new faces and a new colour	Time delay between training task and affective priming task	Facial expression
1	–	Push/pull	X	300	–	–	Neutral
2	X	Push/pull	X	200	–	–	Neutral
3	X	Push/pull	X	200	–	X	Neutral
4a	X	Push/pull	X	200	X	–	Neutral
4b	X	Push/pull	X	200	X	–	Neutral
Woud et al. (2008)	–	Push/pull	X	300	–	–	Neutral
Woud et al. (2011)	–	Push/pull	X	300	–	–	Neutral/angry/ smiling
Woud et al. (2011 2a)	–	Push/pull	–	300	–	–	Neutral
Woud et al. (2011 2b)	–	Left/right	X	300	–	–	Neutral

Apparatus and materials. Ten pictures of neutral faces (five men and five women), selected from the picture set used in the Woud et al. (2008) study, served as training stimuli in the training task and as primes in the affective priming task. For the training task, we created a version of the pictures in which a subtle brown or red filter was placed on the picture using image-processing software. We also used the same twelve targets as Woud et al., namely six adjectives related positively (i.e., pretty, pleasant, nice, friendly, sympathetic, loving) and six adjective related negatively (i.e., unkind, rigorous, irritable, surly, unsympathetic, unfriendly) to sympathy.

The training task was programmed using Visual Basic and was provided to us by one of the authors of Woud et al. (2008). We used a Logitech Attack 3 joystick during this phase. The affective priming task and the rating phase were programmed using the Inquisit software package (Millisecond Software, 2007). The experiment was run on a Dell Dimension 5000 desktop computer with a 17-inch colour monitor that had a refresh rate of 85 Hz.

Procedure. Participants were tested individually in a dimly lit room. After a short introduction about the procedure and the objectives of the experiment, they signed the informed consent form and were seated approximately 50 cm from the computer screen.

During the training task, pictures of faces with neutral expressions were presented in red or brown in the middle of the screen. Participants were instructed to push or pull the joystick as quickly as possible depending on the colour of the presented face. Half of the participants was told to push the brown faces and to pull the red faces; the other half was given the reversed instruction. To resolve the potential ambiguity of the push and pull movements, the pictures became smaller while pushing and larger while pulling the joystick. We told participants that this zoom effect was intended to create the impression that they were pushing the picture away from themselves or pulling the picture towards themselves. Moreover, the zoom effect was contingent upon the action of the participant. Hence, it is likely that participants actually did encode the push response as activate avoidance (i.e., "I increase the distance between myself and the face") and the pull response as activate approach (i.e., "I decrease the distance between myself and the face"). The stimuli disappeared when the joystick was pushed or pulled by 30°. Each trial had to be initiated by the participant by bringing the joystick back into the neutral position and pushing the fire trigger.

The training task consisted of four phases. First, participants completed a practice block of

16 trials with one male and one female face, each presented eight times in red or brown. Then, three training blocks of 96 trials followed with eight experimental faces (four men and four women). Per gender, two faces were always presented in red and two were always presented in brown. During the practice block and the training blocks, the pictures disappeared only when the joystick was moved in the correct direction. Immediately after the third training block, a test block of 16 trials followed in which each experimental face was presented twice, once in red and once in brown. During this block no feedback was given and the pictures disappeared as soon as the joystick was moved 30° from the starting point, irrespective of whether the direction of the move was correct. A final training block of 96 trials was presented after the test block to rule out possible effects of the test block. In total, the training task (excluding practice and test blocks) thus consisted of 384 trials across which each experimental face was pushed or pulled 48 times (i.e., not counting the two combinations in the test block). This is slightly less than the 50 face–movement combinations that produced the biggest effects in the study of Woud et al. (2008). We did not present additional face–movement combinations for 10 or 20 times because the effect of these additional combinations in the Woud et al. study was small.

Whereas Woud et al. (2008) counterbalanced the order of the affective priming task and the rating task, all our participants performed the affective priming task before giving the sympathy ratings. We deviated from the method of Woud et al. in this respect because we were interested primarily in the effects of training on implicit evaluations and because we wanted to avoid an impact of ratings on implicit evaluations. 192 trials were created by randomly combining primes and targets with the restriction that congruent trials (i.e., pulled face and positive target or pushed face and negative target) and incongruent trials (i.e., pulled face and negative target or pushed face and positve target) were presented equally often. The primes were presented in greyscale, that is, without the red or brown filter.

Each trial started with the presentation of the prime for 300 ms, immediately followed by the target. The target remained on the screen until participants gave a response, with a response deadline of 3000 ms. An error message (a red cross) appeared on the screen when participants gave an incorrect response or when they reacted after the response deadline. Participants were required to respond as quickly and as accurately as possible to the category (positive or negative) of the targets by pressing either the "q" or the "m" key on an AZERTY keyboard. The assignment of the response keys to either positive or negative words was counterbalanced across participants.

After the priming task, the explicit evaluation of the faces was measured on a 9-point Likert scale ranging from 0 (*not at all*) to 8 (*very much*). Participants were asked to indicate how sympathetic they found the presented face by clicking on the appropriate box. Faces were presented in greyscale. The picture and the rating scale remained on the screen until a response was given.

Results

Training task. Following Woud et al., we first checked whether training had the desired effect on performance in the training task by analysing the median reaction times (RTs) of the test block. Trials with incorrect responses (1.5% of all trials) were not taken into account. We performed a repeated-measures analysis of variance (ANOVA) with movement as a within-subject variable. The movement could either be congruent (e.g., pull a face that also had to be pulled during the rest of the training task) or incongruent (e.g., pull a face that otherwise had to be pushed away) with the contingencies that were present during the other blocks of the training. The analysis failed to show a significant difference between performance on congruent ($M = 657$, $SD = 143$) and incongruent ($M = 687$, $SD = 188$) trials: $F(1, 28) = 2.15$, $p = .15$, $\eta^2 = .07$.

Affective priming task. RTs on trials with an incorrect response (4.0% of all trials) were not taken into account for the analyses. Median RTs

were analysed using a 2 (Prime: pulled vs. pushed) $\times$ 2 (Target: positive vs. negative) repeated-measures ANOVA. This yielded a significant interaction, $F(1, 28) = 4.27$, $p < .05$, $\eta^2 = .13$. Table 4 shows that the direction of the effect was opposite to that observed by Woud et al. (2008). Overall, RTs on congruent trials (i.e., pulled face and positive target or pushed face and negative target) were slower than on incongruent trials (i.e., pulled face and negative target or pushed face and positive target). A significant main effect of Target indicated that participants were faster on trials with positive targets, $F(1, 28) = 10.74$, $p < .005$, $\eta^2 = .28$. The main effect of Prime was not significant, $F < 1$.

Ratings. For the explicit ratings, data from one participant were missing due to experimenter error. A repeated-measures ANOVA with Picture (pulled vs. pushed) as within-subjects factor did not reveal a significant difference between the ratings of pulled or pushed faces, $F < 1$.

Discussion

Despite following almost exactly the same procedure as Woud et al. (2008), we were unable to replicate their main results. Whereas Woud et al. found that pulled faces facilitated responding to positive targets and pushed faces facilitated responding to negative targets, we found the reversed effect. This contrast effect can be interpreted in a number of ways. First, one could argue that the contrast effect is a training effect nevertheless, thus proving the effectiveness of the training procedure. Moreover, the reversed direction of the effect could be related to the characteristics of the affective priming task rather than to the nature of the implicit evaluation that was induced by the training procedure. For instance, Klauer, Teige-Mocigemba, and Spruyt (2009) recently argued that affective priming effects can shift from assimilative effects (faster when prime and target have the same valence) to contrast effects (faster when the valence of the prime and target differs) as the stimulus onset asynchrony (SOA) increases. The SOA used in our study (300 ms) could be long enough to allow for contrast effects. Although this does not explain why Woud et al. found an assimilative effect despite using the same SOA, this interpretation of the data would imply that the chances of finding an assimilative effect can be increased by decreasing the length of the SOA.

One could also argue, however, that the observed contrast priming effect reflects a genuine reversed learning effect (i.e., pulled faces are more negative than pushed faces). Although this interpretation would still imply that the approach–avoid training method leads to changes in the implicit evaluation of stimuli, it would mean that the direction of the changes is unpredictable. Such unpredictability would seriously undermine the usefulness of the training method.

Finally, one could argue that the contrast effect is a chance finding. This could mean that the approach–avoid training method is generally ineffective and that also the results of Woud et al. were due to chance. It could also mean that the approach–avoid training procedure is effective, but that it did not work in our study. The latter interpretation is in line with the fact that training did not even have a significant effect on performance during the training task. That is, when the contingency between the faces and the movements was temporarily reduced, this did not adversely affect performance during the training task. If there is no effect of training in the training task, it is unlikely that the training effect would generalise to another task such as the affective priming task. On the other hand, the absence of an effect during the test block of the training task might be due to the small number of observations during that block (i.e., 16) and the low power of our statistical tests. From this perspective, it is important to point out that the direction of the effect was, if anything, in the right direction. Also, a strong effect during training does not seem to be a prerequisite for an effect of training on implicit evaluation because Woud et al. (2008) observed only a marginally significant effect during training but a strong effect in the priming task when the faces and movements were combined 20 times.

Because of the results of Experiment 1 do not allow for strong conclusions when considered in isolation, we decided to conduct multiple additional studies. In Experiment 2, we implemented a number of changes that we hoped would increase the training effect and prevent contrast effects during priming.

EXPERIMENT 2

In the approach–avoid training method as introduced by Woud et al. (2008), participants are asked to respond to the colour of the picture. Hence, one cannot be sure that participants process the identity of the faces during training. Because associative learning requires attention (e.g., Mackintosh, 1975), the task to respond to the colour of the faces could reduce the probability that participants learn the contingencies between the (task-irrelevant) identity of the faces and the direction of the movements. We tested this hypothesis by manipulating the extent to which attention was drawn to the identity of the faces. Whereas half of the participants received the same training instructions as in the previous studies (Woud et al., 2008; Experiment 1 of the present paper), the other participants were given the secondary task to give a verbal response whenever one of two target faces was presented on the screen. This secondary task required them to check the identity of each face that was presented during the training task. Furthermore, in order to reduce the probability of contrast effects in the affective priming task, we shortened the SOA.

Method

Participants. Fifty-nine students at Ghent University (48 women, mean age = 20.74, $SD = 1.95$) participated in exchange for course credits or € 8.

Apparatus, materials, and procedure. Participants were randomly assigned to one of two groups. Before starting the training task, participants in the *increased attention group* were given a piece of paper showing two of the eight experimental faces. They were asked to say out loud "A" whenever the first face appeared on the screen and to say "B" whenever the second face was shown. The selection of the two faces and their assignment to the verbal responses varied between participants. Participants in the *standard attention group* performed the training task as in Experiment 1. In the priming task, primes were presented for only 200 ms. In all other respects, Experiment 2 was exactly the same as Experiment 1.

Results

Training task. The data of one participant were excluded from this analysis because the error rate during the test block of the training task was more than three standard deviations (*SD*s) higher than the mean error rate of the total group. RTs on trials with an incorrect response (2.6% of all trials) were also discarded.

A repeated-measures ANOVA with Movement (congruent vs. incongruent) as within-subject factor and Condition (increased attention vs. standard) as between-subject factor revealed a significant interaction, $F(1, 56) = 27.29$, $p < .001$, $\eta^2 = .33$. We performed separate *t*-tests for both groups in order to explore this interaction. In the *increased attention group*, this yielded a significant difference between training congruent ($M = 864$, $SD = 145$) and training incongruent ($M = 1084$, $SD = 239$) trials, $t(27) = 7.39$, $p < .001$. This difference indicates that training was effective and that participants learned which faces they were instructed to push and which to pull. There was no significant effect of Movement in the *standard attention group*: participants did not perform significantly faster on training congruent ($M = 652$, $SD = 682$) trials compared to training incongruent ($M = 682$, $SD = 145$) trials, $t(29) = 1.45$, $p = .16.$, $F < 2.10$, $p > .15$.

Affective priming task. The data of one participant were discarded because the error rate deviated more than three *SD*s from the mean error rate of the total group. In this and all following analyses of priming data, we also removed the data

of participants who were excluded from the analysis of the training data. Because it is likely that these participants did not receive optimal training (e.g., because of many incorrect responses), their data could counteract a possible effect of training on implicit evaluation. None of our conclusions, however, depend on excluding the data of this small number of participants. Finally, RTs on trials with an incorrect response were discarded (6.0% of all trials).

A 2 (Prime: pulled vs. pushed) $\times$ 2 (Target: positive vs. negative) $\times$ 2 (Attention: increased attention group vs. standard attention group) repeated-measures ANOVA yielded a significant three-way interaction, $F(1, 55) = 4.11$, $p < .05$, $\eta^2 = .07$. To follow up on this interaction, separate Prime $\times$ Target ANOVAs were performed for the two groups (see Table 2). In the *increased attention group* the expected Prime $\times$ Target interaction was significant, $F(1, 26) = 7.76$, $p = .01$, $\eta^2 = .23$. Main effects of Prime or Target were not significant, all Fs < 1.66, all ps $> .20$. In the *standard attention condition*, the Prime $\times$ Target interaction was not significant, $F < 1$. A significant main effect of Target was found, $F(1, 29) = 10.75$, $p < .005$, $\eta^2 = .27$, indicating that participants were faster on trials with positive targets. The main effect of Prime was not significant, $F < 1$.

Rating. We performed a 2 (Picture: pulled vs. pushed) $\times$ 2 (Attention: increased attention group vs. standard attention group) repeated-measures ANOVA. A significant main effect of Picture, $F(1, 56) = 5.07$, $p < .05$, $\eta^2 = .08$, showed that pulled faces ($M = 4.82$, $SD = 0.85$) were rated as

being more sympathetic than pushed faces ($M = 4.39$, $SD = 1.10$). The interaction effect was not significant, $F < 1$.

Discussion

The data from Experiment 2 revealed a learning effect but only in the condition with a secondary task that increased attention to the identity of the faces. The attentional manipulation during the training task increased not only the learning effect that was observed during the test trials of the training task but also the learning effect that was observed in the affective priming task. Importantly, in the increased attention group, we were able to induce implicit evaluations as the result of the approach–avoid training task. It is still unclear why we did not find a learning effect in the standard group because this group received the same kind of training then the participants in the original study of Woud et al. (2008) and the participants in our first experiment. The fact that this standard training procedure has resulted in an assimilative effect (i.e., pulled faces liked more than pushed faces; Woud et al., 2008), a reversed effect (Experiment 1), and no effect (Experiment 2), lends support to the conclusion that this procedure is not effective in changing implicit evaluations. Nevertheless, the results of the increased attention group do support the hypothesis that a modified version of the approach–avoid training method can be used to induce implicit evaluations. Note that across groups, we observed for the first time that training resulted in a change in explicit evaluations. This effect was not present in the study of Woud et al.

Table 2. *Mean RTs (in ms) and SDs as a function of prime, target, and condition in Experiment 2*

| | Prime pulled | | | | Prime pushed | | | | |
| | Target positive | | Target negative | | Target positive | | Target negative | | |
Condition	M	(SD)	M	(SD)	M	(SD)	M	(SD)	Prime $\times$ Target interaction
Standard attention	532	(57)	546	(55)	533	(48)	547	(61)	$F(1, 29) = 0.00$, $p = .99$, $\eta^2 = .00$
Increased attention	548	(76)	569	(81)	564	(93)	559	(81)	$F(1, 26) = 7.76$, $p = .01$, $\eta^2 = .23$

EXPERIMENT 3

In an attempt to replicate and extend the promising findings of Experiment 2, we conducted a third experiment. In addition to manipulating attention to the identity of the faces, we varied the time between the training task and the affective priming task to check whether the training effect becomes weaker or stronger after a short time delay. If the approach–avoid training method is to become useful for controlling implicit evaluations, it is important that the effects of the training do not dissipate quickly.

Method

Participants. Eighty students (73 women, mean age = 19.44, $SD = 1.11$) at Ghent University participated in return for course credits or € 8.

Apparatus, materials, and procedure. Participants were randomly allocated to one of four groups. The attention manipulation was identical to that of Experiment 2. Orthogonal to this manipulation, half of the participants performed the affective priming task (SOA = 200 ms) immediately after the training task whereas the other half performed the priming task 20 minutes after finishing the training task. The latter participants spent this waiting period in an adjacent room without any possible distraction.

Results

Training task. The data of three participants had to be discarded: One participant was not able to distinguish between the two colours in which the faces were presented, one participant had an error rate that deviated more than three *SD*s from the mean error rate of the total group mean, and for one participant data were missing due to a technical error. RTs on trials with an error were also discarded (2.4%). We performed a repeated-measures ANOVA with Movement (congruent vs. incongruent) as within-subject factor and

Attention (increased attention group vs. standard attention group) and Delay (delayed vs. immediate affective priming test) as between-subject factors. This analysis yielded a significant Movement × Attention interaction, $F(1, 73) = 8.00$, $p < .01$, $\eta^2 = .10$. Separate *t*-tests showed that in the *increased attention group*, participants responded faster on training congruent ($M = 916$, $SD = 141$) compared to training incongruent ($M = 1113$, $SD = 206$) trials, $t(39) = 6.99$, $p < .001$. In the *standard attention group*, participants also responded more quickly on training congruent ($M = 695$, $SD = 131$) than on training incongruent ($M = 787$, $SD = 218$) trials, $t(36) = 3.92$, $p < .001$, but this effect was smaller ($d = 1.87$ vs. $d = 0.76$, respectively). The ANOVA also showed a main effect of Movement, $F(1, 73) = 59.73$, $p < .001$, $\eta^2 = .45$. In line with the fact that the delay manipulation occurred only after the training task, this factor did not had any effect, all $Fs < 1$.

Affective priming task. The data of two participants were excluded from further analyses because their error rates deviated more than three *SD*s from the mean of the total group. RTs on trials with an incorrect response were also discarded (3.5% of all trials). RTs were analysed using a 2 (Prime: pulled vs. pushed) × 2 (Target: positive vs. negative) × 2 (Attention: with attention vs. without attention) × 2 (Time Delay: with time delay vs. without time delay) ANOVA with repeated measures on the first variable. This analysis revealed a significant main effect of Target, $F(1, 71) = 12.47$, $p = .001$, $\eta^2 = .15$. Participants' RTs were shorter on trials with positive targets. However, neither the interaction between Prime and Target, nor the higher-order interactions involving Prime and Targets reached significance, $Fs < 1$, for these and the other effects. Additional a priori Prime × Target ANOVAs for each condition separately consistently failed to reveal an interaction (see Table 3).

Rating. Data from two participants were missing due to a technical error. A 2 (Picture: pulled

Table 3. *Mean RTs (in ms) and SDs as a function of prime, target, and condition in Experiment 3*

| | Prime pulled | | | | Prime pushed | | | | |
| | Target positive | | Target negative | | Target positive | | Target negative | | |
Condition	M	(SD)	M	(SD)	M	(SD)	M	(SD)	Prime × Target interaction
Standard delayed	578	(60)	597	(65)	578	(48)	595	(56)	$F(1, 17) = 0.06$, $p = .82$, $\eta^2 = .00$
Standard immediate	582	(75)	589	(76)	582	(80)	593	(80)	$F(1, 16) = 0.17$, $p = .68$, $\eta^2 = -.01$
Increased delayed	570	(48)	573	(51)	564	(36)	577	(42)	$F(1, 19) = 2.20$, $p = .16$, $\eta^2 = -.10$
Increased immediate	604	(96)	614	(97)	598	(89)	609	(82)	$F(1, 19) = 0.03$, $p = .88$, $\eta^2 = .00$

Note: A negative effect size (η^2) value indicates that the direction of the priming effect was opposite to that reported by Woud et al. (2008).

vs. pulled) × 2 (Attention: with attention vs. without attention) × 2 (Time Delay: with time delay vs. without time delay) repeated-measures ANOVA did not reveal a main effect of Picture, $F(1, 69) = 0.93$, $p = .34$, $\eta^2 = .01$, nor any interaction involving this variable, all Fs < 2.31, all ps > .13.

Discussion

As in Experiment 2, the attentional manipulation influenced the training effect in the test block of the training task, thus showing that the manipulation was effective and that participants did not simply ignore the instructions of the attention task. Nevertheless, training did not have an effect on implicit evaluations as measured by the affective priming task, neither in the increased attention groups, nor in the standard attention group. The delay manipulation did not result in any effect at all. The effect did not even approach significance in the group that was identical to the increased attention group of Experiment 2 despite the fact that a significant effect of training on implicit evaluation was found in that latter group (see Table 3).

EXPERIMENTS 4A AND 4B

Because of the discordant results of Experiments 2 and 3, we were still unsure about whether we could induce implicit evaluations using the approach–avoid training method. We thus decided to continue our quest. Given that we had found an effect of training on implicit evaluations only when participants had the secondary task of attending the identity of the faces (i.e., increased attention group of Experiment 2), we decided to give this task also to all of the participants in Experiments 4a and 4b.

In addition, we tried to optimise the effect by increasing the similarity between the face stimuli of the training task and those used in the affective priming task. In previous studies, the faces were coloured brown or red during the training task but were presented in greyscale during the affective priming task. Because the transfer of learning from the training to the priming task could depend on the similarity of the stimuli used during training and those used during priming, we manipulated this similarity in Experiments 4a and 4b. Moreover, it is possible that training leads to a change in implicit evaluations of the colours that were presented in that task. For instance, if participants are asked to push red faces away and to pull brown faces towards them, it could be that they develop an implicit preference for red faces over brown faces, irrespective of the identity of the faces. Such learning could overshadow the learning of the relation between the identity of the faces and the direction of the movement. In order to test these hypotheses, we presented three versions of each face during the priming task: One with a red filter, one with a brown filter, and one in greyscale. We also used new faces as primes, that is, faces that were not presented

during the training task. These faces were also presented in brown, red, and greyscale. This allowed us to test whether new faces presented in the pulled colour would elicit more positive implicit evaluations than new faces presented in the pushed colour. However, because of a programming error in Experiment 4a, the data with new faces could only be analysed for Experiment 4b.

Method

Participants. Thirty-eight students (28 women, mean age = 19.02, $SD = 0.96$) at Ghent University participated in Experiment 4a. Fifty-eight other Ghent University students (45 women, mean age = 20.12, $SD = 2.11$) participated in Experiment 4b. They received course credits or € 8 for their participation

Apparatus, materials, and procedure. During the training task, all participants were given the secondary task of detecting two target faces. Hence, the training procedure was identical to that of the increased attention groups of Experiments 2 and 3.

The affective priming task differed in a number of ways from the one used in the previous studies. Most importantly, the eight training faces were presented an equal number of times in brown, red, and greyscale. We also presented four new faces as primes, that is, faces that were not presented during the training task. These faces were taken from the set of faces used by Woud et al. (2008) and were presented an equal number of times in brown, red, and greyscale. However, because of a programming error, in Experiment 4a, new faces were never presented in red before a negative target. Nine positive adjectives and nine negative adjectives that were all related to sympathy were used as targets. 360 trials were created by randomly combining the primes and targets with the restriction that congruent and incongruent trials were presented equally often. The SOA was set at 200 ms. During the rating task, we pre-

sented both the eight training faces and the four new faces in greyscale.

Results

Training task. In both Experiments 4a and 4b, the data of one participant were excluded from all further analyses because of an error rate deviating more than three SDs from the mean of the total group. In both experiments, 3.1% of all trials were discarded because of an incorrect response. We performed repeated-measures ANOVA's with Movement (congruent vs. incongruent) as within-subject factor. This ANOVA revealed a significant difference between training congruent and incongruent trials in Experiment 4a ($M = 792$, $SD = 178$, and $M = 912$, $SD = 254$, respectively), $F(1, 36) = 18.69$, $p < .001$, $\eta^2 = .34$, and in Experiment 4b ($M = 786$, $SD = 192$, and $M = 923$, $SD = 303$, respectively), $F(1, 56) = 24.83$, $p < .001$, $\eta^2 = .30$.

Affective priming task. RTs on trials with an error were discarded (3.9% of all trials in Experiment 4a; 4.0% in Experiment 4b). Because of the programming error in Experiment 4a, only trials with trained faces as primes were analysed in that experiment. The 2 (Prime: pulled vs. pushed) × 3 (Colour: pulled, pushed, or new) × 2 (Target: positive vs. negative) repeated-measures ANOVA on the data of these trials did not reveal the crucial interaction between face and target, $F(2, 35) = 1.22$, $p = .16$, $\eta^2 = .11$, nor any other effects, all Fs < 1. In Experiment 4b, this ANOVA also did not reveal the crucial interaction, $F(2, 55) = 1.08$, $p = .13$, $\eta^2 = .16$, all other Fs < 1. In Experiment 4b, RTs of trials with new faces were analysed using a 3 (Colour: pulled, pushed or new) × 2 (Target: positive vs. negative) repeated-measures ANOVA. This analysis did not yield an interaction, $F(2, 55) = 1.48$, $p = .14$, $\eta^2 = .15$, or main effects, all Fs < 1.

Rating. In both experiments, a repeated-measures ANOVA with Face (pulled, pushed or new) as within-subject factor yielded a significant effect of Face, $F(2, 35) = 6.22$, $p < .01$, $\eta^2 = .26$, for

Experiment 4a, and $F(2, 55) = 4.12$, $p < .05$, $\eta^2 = .13$, for Experiment 4b. Follow-up comparisons showed a significant difference between pulled faces ($M = 4.56$, $SD = 0.86$, in Experiment 4a; $M = 4.62$, $SD = 0.98$, in Experiment 4b) and new faces ($M = 3.82$, $SD = 1.13$, Experiment 4a; $M = 4.19$, $SD = 1.07$, in Experiment 4b), $t(36) = 3.21$, $p < .005$, $t(56) = 2.60$, $p < .05$, for Experiments 4a and 4b, respectively. In Experiment 4a, the difference between pushed faces ($M = 4.50$, $SD = 0.85$) and new faces ($M = 3.82$, $SD = 1.13$) was also significant, $t(36) = 3.06$, $p < .005$. In Experiment 4b, the difference between the mean rating of pushed faces ($M = 4.57$, $SD = 1.04$) and new faces ($M = 4.19$, $SD = 1.07$) was only marginally significant, $t(56) = 1.98$, $p = .07$. However, in neither experiment did we observe a difference between the mean ratings of pushed and pulled faces, ts < .36.

Discussion

The results of Experiments 4a and 4b are highly consistent. Analyses again failed to reveal an effect of training on implicit evaluation of the faces. This failure occurred independently of the visual similarity between the faces presented during the training task and the affective priming task. There was also no evidence for the hypothesis that the colours acquired an implicit valence during the training task. Finally, in the rating task, participants liked both pushed and pulled faces better than faces that were not presented during training. This could be due to the fact that participants had seen the pushed and pulled faces more often (see Zajonc, 1968). Whether a face was repeatedly combined with a push or a pull movement, however, did not have an effect on either implicit or explicit evaluations. These null effects occurred even though performance during the test block of the training task clearly showed that participants had learned the face–movement contingencies.

OTHER STUDIES

The negative results that we obtained in our studies prompted us to examine in more detail the evidence that was obtained outside of our lab. By analysing all the data that were available to us at that time (April 2010), we hoped to obtain a reliable estimate of the size of the effect of this training task on implicit evaluations of novel faces. We therefore asked the authors of Woud et al. to provide us with the raw data of their study and any related studies they might have conducted. We were fortunate to receive the raw data of the seminal study described in Woud et al. (2008) as well as the raw data of five new, as yet unpublished, studies (Woud, Becker, & Rinck, 2011, Experiments 1a, 1b, 1c, 2a, and 2b). The authors also provided us with the aggregated data set of each study, that is, the medians for each participant and each cell of the design as they were entered into the ANOVA reported by Woud and colleagues. In this section, we describe the results of our reanalyses of these data.

Woud et al. (2008)

With regard to the data of the training task and explicit ratings, we were able to reproduce the means and effects reported by Woud et al. (2008). We were also able to reproduce the effects in the affective priming data, provided that we used the aggregated data set of Woud and colleagues and included gender as a variable in the ANOVA. Based on the raw data, we were able recreate all but 75 (4.50%) of the 1,680 data points in the aggregated data set of Woud and colleagues. Despite the small number of deviations, the ANOVA performed on our aggregated data set failed to reproduce the crucial Prime × Target interaction, $F(1, 69) = 1.86$, $p = .18$, $\eta^2 = .03$, as well as the three-way interaction between Prime, Target, and Number of Pairings, $F(2, 69) = 0.13$, $p = .88$, $\eta^2 = .004$.[2] Woud and colleagues recently

---

[2] Details about the nature of these differences can be obtained upon request. Note that the third author of the Woud et al. (2008) paper confirmed that he alone had performed the original analyses and that only he was responsible for any errors in these analyses.

Table 4. *Mean RTs (in ms) and SDs as a function of prime, target, and experiment*

	Prime pulled				Prime pushed				
	Target positive		Target negative		Target positive		Target negative		
Experiment	M	(SD)	M	(SD)	M	(SD)	M	(SD)	Prime × Target interaction
1	528	(55)	540	(51)	523	(55)	547	(56)	$F(1, 28) = 4.27$, $p < .05$, $\eta^2 = -.13$
2	539	(67)	557	(69)	548	(74)	553	(71)	$F(1, 56) = 3.51$, $p = .07$, $\eta^2 = .06$
3	584	(72)	593	(74)	581	(67)	594	(67)	$F(1, 74) = 0.80$, $p = .37$, $\eta^2 = -.01$
4a	587	(99)	584	(84)	590	(83)	592	(90)	$F(1, 36) = 0.20$, $p = .66$, $\eta^2 = -.01$
4b	589	(87)	598	(91)	596	(77)	590	(77)	$F(1, 56) = 2.20$, $p = .14$, $\eta^2 = .04$
Woud et al. (2008)	581	(76)	600	(75)	579	(74)	592	(72)	$F(1, 69) = 1.66$, $p = .20$, $\eta^2 = .02$
Woud et al. (2011 1a)	565	(90)	598	(113)	569	(91)	585	(82)	$F(1, 28) = 4.23$, $p = .04$, $\eta^2 = .05$
Woud et al. (2011 1b)	566	(61)	586	(65)	561	(65)	583	(66)	$F(1, 81) = 0.07$, $p = .80$, $\eta^2 = .01$
Woud et al. (2011 1c)	586	(109)	606	(120)	573	(102)	608	(135)	$F(1, 82) = 1.68$, $p = .20$, $\eta^2 = .02$
Woud et al. (2011 2a)	571	(73)	594	(75)	572	(74)	590	(67)	$F(1, 97) = 0.75$, $p = .39$, $\eta^2 = .01$
Woud et al. (2011 2b)	566	(107)	598	(104)	568	(88)	601	(110)	$F(1, 97) = 0.02$, $p = .88$, $\eta^2 = .00$
Overall	571	(86)	591	(90)	571	(81)	588	(88)	$F(1, 768) = 0.86$, $p = .35$, $\eta^2 = .00$

Note: A negative effect size (η^2) value indicates that the direction of the priming effect was opposite to that reported by Woud et al. (2008).

published a correction note in which they acknowledged that the effects reported by Woud et al. (2008) cannot be reproduced in a reanalysis of the raw data. In the correction note, they reported the results of an ANOVA that did not include gender of the prime faces as a variable. This ANOVA revealed a marginally significant interaction between Prime and Target, $F(1, 69) = 3.91$, $p = .052$, $\eta^2 = .05$, but not a three-way interaction between Prime, Target, and Number of Pairings, $F < 1$. We could indeed reproduce the Prime × Target interaction effect when using this type of ANOVA, but not when gender was included as a variable (see above) or when number of pairings was not included as a variable (see Table 4).

Woud et al. (2011)

We also obtained the data of another set of studies that have been described in a recent unpublished paper (Woud et al., 2011). Because we were able to recreate all the effects reported in this new paper, we provide only a brief summary of the

various studies (see Table 1 for an overview) and results (see Table 4 for an overview of the priming effects) as they were available to us in April 2010.[3]

In Experiments 1a, 1b, and 1c of Woud et al. (2011), the faces used in the training task had a neutral, angry or smiling expression, respectively. Each face was pushed or pulled 50 times. In all other respects, the experiment was identical to that of Woud et al. (2008). During the test block of the training task, the expected learning effect (faster on movement congruent trials) occurred in each of the experiments but most clearly so in Experiment 1c (smiling faces) than in Experiments 1a (neutral faces) and 1b (angry faces). The expected priming effect, however, reached significance only in Experiment 1a (see Table 4). No learning effects were found in the rating data.

Experiment 2a of Woud et al. (2011) was identical to the original study of Woud et al. (2008) except that the size of the pictures during the training phase did not change when pulling or pushing the joystick. Again the expected learning effect was found in the training data. There was,

---

[3] All results are based on the data that were available to us in April 2010, except for the results of Experiment 1a because we were informed about an error in this dataset only after the submission of this paper in April 2011.

however, no trace of a learning effect in the priming data (see Table 4).

Woud et al. (2011) also reported a control experiment (Experiment 2b) in which participants did not pull or push the joystick during the learning phase. Instead, they were asked to push away a picture (as indicated by the fact that a picture became smaller) or pull the picture towards them (as indicated by the fact that a picture became larger) by moving the joystick to the left or the right. Although left or right movements were not intrinsically related to approach or avoidance, they can be regarded as approach and avoidance responses because they were linked via instructions with approach and avoidance (see De Houwer, Crombez, Baeyens, & Hermans, 2001; Eder & Rothermund, 2008, for evidence supporting this assumption). As Table 4 shows, however, training with these responses did not result in a subsequent affective priming effect.

Overall analyses

In order to obtain a reliable estimate of the size of the effect of the approach–avoid training that was devised by Woud et al. (2008), we aggregated the data of all experiments, resulting in a total sample of 769 (568 for the explicit ratings) participants. Because the design and procedure of the various experiments was not entirely identical, we included experiment as a between-subjects factor. A repeated-measures ANOVA with Movement (congruent vs. incongruent) as within-subjects factor and Experiment (1–9) as between-subjects factor revealed a strong learning effect in the test block of the training task. RTs were shorter on training congruent ($M = 685$, $SD = 157$) compared to training incongruent ($M = 766$, $SD = 246$) trials, $F(1, 758) = 214.60$, $p < .001$, $\eta^2 = .22$. Neither the main effect of Experiment, $F(10, 758) = 2.15$, $p = .09$, $\eta^2 = .22$, nor the interaction, $F(10, 758) = 1.67$, $p = .12$, $\eta^2 = .18$, reached significance.

We analysed the affective priming data using a Prime (pulled vs. pushed) × Target (positive vs. negative) × Experiment (1–9) mixed ANOVA with repeated measures on the first two variables.

For Experiments 4a and 4b, only trials with trained faces and trained colours were analysed. Most importantly, neither the interaction between prime and target, $F(1, 758) = 0.40$, $p = .53$, $\eta^2 = .00$, nor the interaction between Prime, Target, and Experiment, $F(10, 758) = 1.58$, $p = .11$, $\eta^2 = .02$, reached significance. The ANOVA revealed only main effects of Target, $F(1, 758) = 132.67$, $p < .001$, $\eta^2 = .15$, Experiment, $F(10, 758) = 2.42$, $p = .008$, $\eta^2 = .03$, and an interaction between Target and Experiment, $F(10, 758) = 4.90$, $p < .001$, $\eta^2 = .06$ (all other effects $F < 1$). Importantly, our statistical test of the crucial Prime × Target interaction had maximal power (1.00) to reveal a small effect (Cohen's $d = 0.20$; Faul, Erdfelder, Lang, & Buchner, 2007).

Explicit ratings were not registered in Experiments 2a and 2b of Woud et al. (2011). A repeated-measures ANOVA with Picture (pulled vs. pushed) as within-subject and Experiment (1–7) as between-subjects factor showed no difference between the ratings of pulled and pushed faces, all $Fs < 1$.

Analyses of affective priming data per type of procedure

Although the overall analysis has the advantage of increased power because of the large number of data points, it is not optimal for revealing the effects of different variants of the approach–avoid training procedure. We therefore conducted additional analyses that were based on a distinction between different types of training procedures. If certain variants of the procedure do produce reliable effects of training on implicit evaluations, this would demonstrate not only that the effects can be found but would also provide important information about the boundary conditions of those effects.

We identified two homogeneous sets of procedures and one miscellaneous set. The first type of procedure corresponds to the procedure used in the original study of Woud et al. (2008) and (with minor changes; see above) in Experiment 1 ($N = 29$), the standard condition of Experiment

2 ($N = 30$), the standard condition with immediate test of Experiment 3 ($N = 17$), Woud et al. (2008; $N = 70$) and Woud et al. (2011, Experiment 1a, $N = 83$). Hence, we have the data of 229 participants who experienced this procedure. The second type concerns the increased attention procedure that generated a significant effect in our second experiment. Apart from the increased attention condition of Experiment 2 ($N = 27$), this procedure was used also in the increased attention condition with immediate test of Experiment 3 ($N = 20$), Experiment 4a (only test trials with trained faces and colours; $N = 36$) and Experiment 4b (only test trials with trained faces and trained colours; $N = 57$), leading to a total sample of 141 participants. The third type concerns procedures that deviate from the other two. It encompasses the procedures used in the standard condition with increased delay of Experiment 3 ($N = 18$), the increased attention condition with increased delay of Experiment 3 ($N = 20$), Woud et al., 2011, Experiments 1b ($N = 82$), 1c ($N = 83$), 2a ($N = 98$), and 2b ($N = 98$), resulting in a total sample of 399 participants.

For each type of procedure, we analysed the affective priming data using a Prime (pushed or pulled) $\times$ Target (positive or negative) ANOVA. None of the analyses revealed the crucial Prime $\times$ Target interaction, $F(1, 228) = 2.92$, $p = .10$, $\eta^2 = .01$, for the first type, $F(1, 140) = 2.39$, $p = .12$, $\eta^2 = .02$, for the second type, and $F(1, 398) < 1$, $p = .47$, $\eta^2 = .00$, for the third type. Nevertheless, all analyses had acceptable power (0.85, 0.65, and 0.98, respectively) to detect a small effect (Cohen's $d = 0.20$),

and maximal power (1.00) to detect a median sized (Cohen's $d = 0.50$) and large effect (Cohen's $d = 0.80$; Faul et al., 2007). Finally, an overall analysis with Prime, Target, and Procedure Type as variables failed to reveal a three-way interaction, $F(1, 766) = 2.13$, $p = .12$, $\eta^2 = .00$. Hence, we did not find clear evidence for the hypothesis that the learning effect in the priming data depended on the type of procedure.[4]

GENERAL DISCUSSION

We conducted five experiments that were designed to replicate and extend the findings of Woud et al. (2008). Three of those experiments consisted of multiple between-subjects conditions. The expected effect of approach–avoid training on implicit evaluations (more positive implicit evaluations towards pulled faces than pushed faces) was observed in only one of those conditions (the increased attention condition of Experiment 2). In Experiment 1, we observed a significant reversed effect (more positive implicit evaluations of pushed faces than pulled faces). In the other seven conditions (standard attention condition of Experiment 2, all four conditions of Experiment 3, Experiment 4a, and Experiment 4b), no effect of training on implicit evaluations was observed.

These discouraging results prompted us to reanalyse the data of Woud et al. (2008). Importantly, we failed to reproduce the effect of training on implicit evaluation that was reported their paper. In a correction note, the authors of Woud et al. (2008) acknowledged that they also could not reproduce the effect that they originally

---

[4]We repeated all the analyses reported in Table 4 and all the analyses per type of procedure using the mean RTs during the priming phase as the dependent variable. For each participant, we first calculated the mean and standard deviation of all RTs on trials with a correct response. Next, RTs that deviated more than 2.5 SDs from the participant's mean were discarded. Finally, mean RTs for each combination of prime and target were calculated based on the remaining RTs within the corresponding cell. These analyses led to the same conclusions as the analyses of the median RTs except for the analysis of the mean RT data of Experiment 4a, which revealed a significant Prime $\times$ Target interaction, $F(1, 36) = 4.35$, $p = .04$. However, contrary to expectations, mean RTs were longer on congruent trials (i.e., $M = 600$, $SD = 91$, for trials with pulled primes and positive targets, and $M = 614$, $SD = 97$, for trials with pushed primes and negative targets) than on incongruent trials (i.e., $M = 606$, $SD = 90$, for trials with pushed primes and positive targets, and $M = 594$, $SD = 89$, for trials with pulled primes and negative targets). Finally, for all experiments, we analysed the error data as well. The pattern of results was virtually the same and led to the same conclusions as the analyses of the (median and mean) RTs. There was no sign of a speed–accuracy trade-off.

reported although they did still find a tendency for an effect in one type of analysis. In a more recent paper, Woud et al. (2011) observed a significant effect of training on implicit evaluations in one of five conditions (Experiment 1a) but no effect in four other conditions (Experiments 1b, 1c, 2a, and 2b).

Based on this divergent pattern of results across the various individual studies and conditions, it is difficult to draw firm conclusions, especially because none of the studies had sufficient power to detect a small effect or even a medium-sized effect (given that sufficient is defined as an 80% chance of detecting an effect with Cohen's $d = 0.20$ or 0.50 when using a significance level of $p = .05$ and a two-sided test (see Faul et al., 2007). In order to alleviate the problems of limited power, we conducted an overall analysis that included all the data that we collected or that were provided to us by Woud and colleagues. Despite the fact that this overall analysis had maximal power to detect even a small effect, the crucial priming effect did not emerge, hence failing to provide evidence for the effect of approach–avoid training on the implicit evaluations that are evoked by novel faces. The results of the overall analysis are important because they argue against the hypothesis that the observed null effects are due to a lack of power. Moreover, because the overall analysis included data from several variations of the procedure of Woud et al. (2008), it provides important information about the specificity of the effects that are produced by this procedure (see Fiedler, 2011, for a detailed discussion of the merits of testing effects under a variety of conditions). Although it might be the case that some specific instantiations of the procedure do produce reliable effects, the fact that effects do not emerge across a range of (relatively minor) variations of this procedure suggests that those effects are subjected to subtle but important boundary conditions and moderators.

Nevertheless, it is also important to try to identify those conditions under which a procedure does produce an effect. The overall analysis is not optimal to achieve this aim because it aggregates

data that were generated across a variety of conditions. We therefore identified two more homogeneous sets of procedures. The first set included procedures identical or highly similar to the procedure introduced by Woud et al. (2008, condition with 50 pairings). Such a procedure resulted in a (marginally) significant priming effect in the expected direction in the (re)analyses of the data of Woud et al. (2008, 2010, Experiment 1a) but a significant reversed effect in our first experiment and no effect in the standard conditions (without delay) of Experiments 2 and 3. When combining all of these data, we still did not find a significant effect of training on implicit evaluations despite adequate power to detect a small effect. This suggests that the (marginally) significant effects that were observed by Woud et al. (2008, 2011) might have been due to Type I errors (but see below). The second homogeneous set consisted of procedures that aimed to increase attention to the identity of the faces during the training task and that measured implicit evaluations immediately after training. Such a procedure produced a large priming effect in the expected direction in Experiment 2. However, it failed to produce an effect in Experiments 3, 4a, and 4b. To examine whether the null findings were due to a lack of power or whether the significant effect was due to a sampling error, we aggregated all the relevant data from Experiments 2, 3, 4a, and 4b. An analysis of these data did not reveal a significant effect of training on implicit evaluations despite acceptable power to detect a small effect (0.65) and maximal power to detect a medium sized or large effect. Again, the one significant result that we obtained with the increased attention procedure thus appears to be due to a Type I error. Finally, when we combined the data of participants who experienced other variants of the Woud et al. (2008) procedure, there was also no sign of changes in implicit evaluations despite maximal power to detect even a small effect.

When considered as a whole, our data and those provided to us by Woud and colleagues thus provide little evidence for the effect of approach–avoid training on implicit evaluations. Also, they

do not reveal variants of this procedure that are effective. It is important, however, that these findings are not interpreted as conclusive evidence for the ineffectiveness of approach–avoid training. It might well be that approach–avoid training can influence implicit evaluations but that we failed to observe these effect for one of several possible reasons. First, because our tests only had sufficient power to detect small effects, we cannot exclude the possibility that the effect of training implicit evaluations is genuine but smaller than what is considered to be a small effect. Second, and more importantly, some of the null results in our studies and those of Woud and colleagues might have been due to subtle aspects of the procedure that counteract the training effects. It is, for instance, striking that Woud et al. (2008) and Woud et al. (2011, Experiment 1a) both found a small but (marginally) significant effect of training on implicit evaluations in the expected direction. In Experiments 1, 2 (standard condition), and 3 (standard condition with immediate test), we used a procedure that was highly similar to the procedure used in the studies of Woud and colleagues but obtained either a reversed (Experiment 1) or no effect (Experiments 2 and 3) in the affective priming task. It is possible that our results deviated from those of Woud and colleagues because of subtle differences in the procedure that we used or the samples that we tested. At present, it is not clear what these differences might be. After the completion of our overall analysis, Woud and colleagues informed us that they are still collecting new data. If those data reveal effects in line with their earlier studies, new research is required to identify those procedural elements that moderate the effects. Regardless of the outcome of those new studies, the large data set that was considered in this paper does allow for the conclusion that there are clear boundary conditions with regard to the effectiveness of the approach–avoid training procedure as introduced by Woud and colleagues. Moreover, we hope that our overview of the currently available data encourages researchers to treat isolated (margin-

ally) significant training effects with caution until they have been proven to be robust and replicable.

Although it is difficult do draw strong conclusions about why we failed to obtain replicable effects of approach–avoid training on implicit evaluations, our data do allow us to examine possible moderators of these effects. First, as noted by an anonymous reviewer, the impact of the training contingencies on implicit evaluations might depend on the extent to which those contingencies influence performance during the test trials of the training phase. If there is no impact of the contingencies on performance during the test phase, there is little reason to expect that they will influence performance during the subsequent priming phase. From this perspective, it is important to note that the training contingencies had a clear impact on performance during the training test phases of Experiments 3, 4a, and 4b. Hence, it is unlikely that lack of effect of training on implicit evaluation in those experiments was due to a failure to register the face-response contingencies during training. Moreover, we found little evidence for a correlation between the magnitude of the training effect during the training phase and during the priming phase. In Experiment 2, a significant positive correlation did emerge, $r = .38$, $p = .003$. In Experiment 1c of Woud et al. (2011), however, a significant negative correlation was found, $r = -.32$, $p = .004$. In the other experiments, the correlation was not significant, $-.20 < r < .23$. When all the data were pooled ($N = 769$), there was a trend for a negative correlation, $r = -.07$, $p = .07$. Hence, our data provide little support for the idea that the effect of training contingencies on implicit evaluations is related to the effectiveness of training as assessed on the test trials of the training phase.

Second, at the end of each of our experiments (i.e., Experiments 1, 2, 3, 4a, 4b), we asked participants to indicate whether they liked or disliked the zoom effect that occurred when pulling a picture towards them. We reasoned that if participants dislike the zoom effect, this might interfere with the transfer of valence from

the responses to the pictures. Across all experiments, 30% of the participants said that they disliked the zoom effect. We then reanalysed the data using ANOVAs with Liking of the Zoom Effect (like or dislike) as between-subjects variable and Prime (pulled or pushed) and Target (positive or negative) as within-subjects variables. Liking of the zoom effect did not moderate priming in Experiments 1, 2, 3, and 4b, as indicated by non-significant three-way interactions, $Fs < 1$. In Experiment 4a, the three-way interaction approached significance, $F(1, 35) = 3.41$, $p = .07$, but only 7 out of 37 participants disliked the zoom effect in that experiment. Moreover, if anything, the priming effect in Experiment 4a was smaller for participants who liked the zoom effect then for the few participants who disliked the zoom effect. Because the analyses of the individual experiments were underpowered, we repeated the ANOVA on a data set that included the data from our five experiments ($N = 255$) and a data set that included only the attention enhancement conditions ($N = 141$). Neither analyses revealed a three-way interaction, $F(1, 253) = 1.39$, $p = .24$, and $F(1, 139) = 1.56$, $p = .21$, respectively. Note, however, that even the latter two analyses did not have sufficient power to detect a small effect (0.31 and 0.18, respectively). They did have acceptable power to detect a medium sized effect (0.95 and 0.75, respectively). In sum, we found little evidence to support the idea that subjective liking of the zoom effect moderated the impact of training on implicit evaluations.

Until now, we have focused on studies that were conducted at our lab or by Woud and colleagues. All of these studies have in common that unfamiliar faces were used as stimuli, colour was relevant during training, an affective priming task was used to capture implicit evaluations, and participants were psychology students. We are aware of a number of other studies that deviate in important ways from the studies considered in this paper but that also examine the effects of approach–avoid training on implicit evaluations. First, Huijding et al. (2009) showed pictures of two different marsupials (a quokka and a quoll) to children unfamiliar with these animals. They asked the children to push pictures of one type of animal away and to pull pictures of the other animal towards them using a joystick. We did not include the results of this study in the overall analysis because it involved children rather than adults and because the identity of the animals (i.e., quokka or quoll) was relevant in this version of the task rather than the colour of the pictures. Although Huijding et al. found an effect of approach–avoid training on explicit evaluations, there was no effect of training on implicit evaluations.

Several other studies examined the effect of approach–avoid training on implicit evaluations that are evoked by highly familiar and ambivalent attitude objects. In studies by Kawakami and colleagues (Kawakami, Phills, Steele, & Dovidio, 2007; Phills, Kawakami, Tabi, & Nadolny, 2011, Experiment 4), participants made approach and avoid responses when presented with pictures of faces of Black or White persons. In some experiments, the correct movement was indicated by the skin colour of the presented face. In other experiments, a different cue determined the to-be-executed response. After the training phase, participants completed a racial Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) designed to assess implicit evaluations of Black and White people in general. The results of four experiments revealed more favourable implicit evaluations of Black versus White people in participants who repeatedly pulled Black faces and pushed White (or Asian) faces away than in participants who received the reversed training task (push away Black faces and pull White faces). In another study, Kawakami, Steele, Cifa, Phills, and Dovidio (2008, Experiment 1) found an effect of training to approach math-related stimuli on implicit evaluations of math, be it only for women who initially did not identify themselves with math.[5]

Based on the evidence provided by Kawakami and colleagues, it seems safe to conclude that

implicit evaluations can be influenced by approach–avoid training procedures. The question remains why the data that were considered in this paper did not provide clear evidence for these effects whereas the studies of Kawakami and colleagues (Kawakami et al., 2007, 2008; Phills et al., 2011) did. Because of the many differences between these two sets of studies, we can only speculate about the possible moderators.

First, the nature of the stimuli might be important. In line with Woud et al. (2008), we used faces of unfamiliar people as stimuli whereas Kawakami et al. used exemplars of highly familiar categories towards which participants probably had strong pre-existing feelings. On the one hand, one would expect that installing new implicit evaluations towards novel stimuli would be easier than changing existing implicit evaluations (see Gregg, Seibt, & Banaji, 2006, for evidence supporting this idea). On the other hand, learning research has established that more learning takes place as the result of events that violate expectations compared to events that are in line with or do not violate expectations (e.g., Rescorla & Wagner, 1972). From this perspective, approaching disliked objects and avoiding liked objects might be coded as events that strongly contradict expectations, thus leading to large changes in preferences. Approaching unknown objects or known liked objects and avoiding unknown objects or known disliked objects might not violate expectations that much and might thus results in fewer changes in preferences. At present, however, this idea remains speculative and requires further testing.

The nature of the stimuli might be important also for other reasons. Implicit evaluations of familiar ambivalent stimuli such as racial groups are known to be malleable. For instance, implicit evaluations of Blacks tend to be less negative when participants think about liked Black athletes than when they think about disliked Black

criminals (e.g., Mitchell, Nosek, & Banaji, 2003). Hence, it is possible that the approach–avoid training effects that were observed by Kawakami et al. (2007; Phills et al., 2011) were due to changes in the accessibility of existing associations in memory rather than instances of additional learning (see Gawronski & Bodenhausen, 2006, for a detailed discussion of this distinction).

A second potential reason for the success of the studies of Kawakami et al. (2007, 2008; Phills et al., 2011) concerns their use of the IAT rather than the affective priming task for capturing implicit evaluations. Affective priming effects tend to be less reliable and smaller in magnitude than IAT effects (see De Houwer, Teige-Mocigemba, Spruyt, & Moors, 2009, for a review), thus reducing the chance of observing effects of approach–avoid training on affective priming effects compared to IAT effects.

Third, the training procedures of Kawakami et al. (2007, 2008; Phills et al., 2011) were focused on changing the evaluation of abstract concepts such as "Blacks" and involved the presentation of many different stimuli that were related to those concepts. The procedure as introduced by Woud et al. (2008), on the other hand, aims to change the evaluation of several individual stimuli. It is possible that the more focused approach of Kawakami et al. is more likely to results in changes in implicit evaluations.

Because of the many differences between the procedure as introduced by Woud et al. (2008) and the procedures used by Kawakami et al. (2007, 2008; Phills et al., 2011), it is difficult to draw strong conclusions about the possible moderators of the effects of approach–avoid training on implicit evaluations. We would like to repeat, however, that our research does show that the effects of approach–avoid training procedures are dependent on important but subtle boundary conditions and moderators. As Fiedler (2011) noted in a recent meta-theoretical paper, "being

---

[5] Wiers and colleagues (Wiers, Eberl, Rinck, Becker, & Lindenmeyer, 2011; Wiers, Rinck, Kordts, Houben, & Strack, 2010) used a training procedure similar to that of Woud et al. (2008) but tested its impact on measures of approach–avoidance behaviour rather than on implicit evaluations. Because it is not clear whether the findings of Wiers and colleagues generalise to implicit evaluations, they will not be discussed further.

explicit about restrictions and crucial catalysts of a phenomenon can be enlightening and conducive to important theoretical insights, rather than being a concession of weakness and invalidity. ... It rather emphasises the value of creative search for subtle boundary conditions and easily overlooked moderators that are only implicit in the restricted samples of stimuli, tasks, and experimental contexts". We hope that our paper provides the impetus for further research on the boundary conditions and moderators of the effects of approach–avoid training on implicit evaluations.

REFERENCES

Baccus, J. R., Baldwin, M. W., & Packer, D. J. (2004). Increasing implicit self-esteem through classical conditioning. *Psychological Science*, *15*, 498–502.

De Houwer, J. (2009). How do people evaluate objects? A brief review. *Social and Personality Psychology Compass*, *3*, 36–48.

De Houwer, J. (2011). Why the cognitive approach in psychology would profit from a functional approach and vice versa. *Perspectives on Psychological Science*, *6*, 202–209.

De Houwer, J., Crombez, G., Baeyens, F., & Hermans, D. (2001). On the generality of the affective Simon effect. *Cognition and Emotion*, *15*, 189–206.

De Houwer, J., & Hermans, D. (2010). Do feelings have a mind of their own? In J. De Houwer & D. Hermans (Eds.), *Cognition and Emotion: Reviews of current research and theories* (pp. 38–65). Hove, UK: Psychology Press.

De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit measures: A normative analysis and review. *Psychological Bulletin*, *135*, 347–368.

Eder, A., & Rothermund, K. (2008). When do motor behaviors (mis)match affective stimuli? An evaluative coding view of approach and avoidance reactions. *Journal of Experimental Psychology: General*, *137*, 262–281.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191.

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology*, *69*, 1013–1027.

Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology*, *54*, 297–327.

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, *50*, 229–238.

Ferguson, M. J., & Zayas, V. (2009). Automatic evaluation. *Current Directions in Psychological Science*, *18*, 362–366.

Fiedler, K. (2011). Voodoo correlations: A severe methodological problem, not only in social neurosciences. *Perspectives on Psychological Science*, *6*, 163–171.

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin*, *132*, 692–731.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, *74*, 1464–1480.

Huijding, J., Field, A. P., De Houwer, J., Vandenbosch, K., Rinck, M., & van Oeveren, M. (2009). A behavioural route to dysfunctional representations: The effects of training approach or avoidance tendencies towards novel animals in children. *Behaviour Research and Therapy*, *47*, 471–477.

Kawakami, K., Phills, C. E., Steele, J. R., & Dovidio, J. F. (2007). (Close) distance makes the heart grow fonder: Improving implicit racial attitudes and interracial interactions through approach behaviors. *Journal of Personality and Social Psychology*, *92*, 957–971.

Kawakami, K., Steele, J. R., Cifa, C., Phills, C. E., & Dovidio, J. F. (2008). Approaching math increases math = me, math = pleasant. *Journal of Experimental Social Psychology*, *44*, 818–825.

Klauer, K. C., Teige-Mocigemba, S., & Spruyt, A. (2009). Contrast effects in spontaneous evaluations:

A psychological account. *Journal of Personality and Social Psychology, 96*, 265–287.

Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review, 82*, 276–298.

Millisecond Software. (2007). *Millisecond Inquisit 2.0* [Computer Software]. Seattle, WA: Millisecond Software.

Mitchell, J. P., Nosek, B. A., & Banaji, M. R. (2003). Contextual variations in implicit evaluation. *Journal of Experimental Psychology: General, 132*, 455–469.

Phills, C. E., Kawakami, K., Tabi, E., & Nadolny, D. (2011). Mind the gap: Increasing associations between the self and Blacks with approach behaviors. *Journal of Personality and Social Psychology, 100*, 197–210.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York, NY: Appleton.

Roefs, A., Huijding, J., Smulders, F. T. Y., de Jong, P. J., Wiers, R. W., & Jansen, A. (2011). The use of indirect measures in psychopathology research: A review. *Psychological Bulletin, 37*, 149–193.

Wiers, R. W., Eberl, C., Rinck, M., Becker, E., & Lindenmeyer, J. (2011). Re-training automatic action tendencies changes alcoholic patients' approach bias for alcohol and improves treatment outcome. *Psychological Science, 22*, 490–497.

Wiers, R. W., Rinck, M., Kordts, R., Houben, K., & Strack, F. (2010). Re-training automatic action-tendencies to approach alcohol in hazardous drinkers. *Addiction, 105*, 279–287.

Wiers, R. W., & Stacy, A. W. (2006). *Handbook of implicit cognition and addiction*. Thousand Oaks, CA: Sage.

Woud, M. L., Becker, E. S., & Rinck, M. (2008). Implicit evaluation bias induced by approach and avoidance. *Cognition and Emotion, 22*, 1187–1197.

Woud, M. L., Becker, E. S., & Rinck, M. (2011). *Does approach–avoidance training affect the evaluation of faces? Findings and boundaries*. Manuscript submitted for publication.

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology, 9*, 1–27.

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35*, 151–175.