The following full text is a publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/202866

Please be advised that this information was generated on 2019-05-13 and may be subject to change.
The processing of task-irrelevant emotion and colour in the Approach-Avoidance Task

Xijia Luo, Mike Rinck, Harold Bekkering & Eni S. Becker

To cite this article: Xijia Luo, Mike Rinck, Harold Bekkering & Eni S. Becker (2019) The processing of task-irrelevant emotion and colour in the Approach-Avoidance Task, Cognition and Emotion, 33:3, 548-562, DOI: 10.1080/02699931.2018.1497951

To link to this article: https://doi.org/10.1080/02699931.2018.1497951

© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

Published online: 09 Jul 2018.

Submit your article to this journal

Article views: 676

View Crossmark data
The processing of task-irrelevant emotion and colour in the Approach-Avoidance Task

Xijia Luo, Mike Rinck, Harold Bekkering and Eni S. Becker

ABSTRACT
When processing information about human faces, we have to integrate different sources of information like skin colour and emotional expression. In 3 experiments, we investigated how these features are processed in a top-down manner when task instructions determine the relevance of features, and in a bottom-up manner when the stimulus features themselves determine process priority. In Experiment 1, participants learned to respond with approach-avoidance movements to faces that presented both emotion and colour features (e.g. happy faces printed in greyscale). For each participant, only one of these two features was task-relevant while the other one could be ignored. In contrast to our predictions, we found better learning of task-irrelevant colour when emotion was task-relevant than vice versa. Experiment 2 showed that the learning of task-irrelevant emotional information was improved in general when participants’ awareness was increased by adding NoGo-trials. Experiment 3 replicated these results for faces and emotional words. We conclude that during the processing of faces, both bottom-up and top-down processes are involved, such that task instructions and feature characteristics play a role. Ecologically significant features like emotions are not necessarily processed with high priority. The findings are discussed in the light of theories of attention and cognitive biases.
processing pathways are interacting with each other in a single task.

Because of the high ecological significance of emotions, it is commonly hypothesized that emotional stimuli are processed with high priority. For instance, studies of attention seem to suggest that the processing of emotional stimuli is involuntary and automatic (see Koster, Crombez, Verschuere, & De Houwer, 2006; Williams, Mathews, & MacLeod, 1996). By using the Approach-Avoidance Task (AAT; Rinck & Becker, 2007), several studies showed that participants’ movements are also influenced by the emotional valence of the stimuli: Approach movements are facilitated by positive, pleasant stimuli (e.g. smiling faces or butterflies), whereas avoidance movements are easier in response to negative, unpleasant, or threatening stimuli (e.g. angry faces or spiders). This pattern was observed in many studies showing that emotional information affects participants’ approach-avoidance behaviours (Krieglmeyer, Deutsch, De Houwer, & De Raedt, 2010; reviews: Laham, Kashima, Dix, & Wheeler, 2015; Phaf, Mohr, Rotteveel, & Wicherts, 2014). The pattern was also observed when participants were instructed to react to non-emotional aspects of the stimuli, such as whether a picture shows a face or a puzzle (Heuer, Rinck, & Becker, 2007), or whether a face picture is printed in greyscale or sepia (Roelofs et al., 2010).

The assumption that emotional facial expressions in particular are processed with high priority is also supported by studies which used the AAT to train approach-avoidance responses rather than to merely measure them. For instance, Taylor and Amir (2012) trained participants to approach pictures of smiling and neutral faces by responding to the colour of a border around the picture. Afterwards, the trained participants displayed stronger approach of smiling faces and more social approach behaviour during a real social interaction, and they elicited more positive reactions from their interaction partners. In a similar study, Rinck et al. (2013) trained socially anxious participants to react to the colour of pictures of emotional faces (approach grey pictures, avoid sepia pictures). For one group of participants, all happy faces were presented in grey-scale, for the other group they were presented in sepia. The connection between colour and emotion was not made explicit. Nevertheless, participants who had been trained to approach grey-happy faces developed a tendency to approach happy faces. Moreover, they also recovered more quickly from a social stress task. These findings are in line with the assumption that emotional information is processed with high priority, even when it is irrelevant for the task at hand.

The latter studies illustrate that the relative priority of emotional vs. physical stimulus features is also relevant in computer-based training procedures, summarized under the name Cognitive Bias Modification (CBM). In many clinical studies, it was found that patients suffering from emotional disorders such as phobias or depression often show cognitive biases when it comes to the processing of emotional stimuli. This has led to the development of CBM training procedures which aim to train healthier processing styles, mainly in the areas of attention, interpretation, memory, and approach-avoidance (Mogoase, David, & Koster, 2014). For instance, using a joystick-based approach-avoidance task, alcohol-addicted patients were successfully trained to automatically avoid rather than approach alcohol-related stimuli, thereby reducing relapse rates even one year later (Eberl et al., 2013; Wiers, Eberl, Rinck, Becker, & Lindenmeyer, 2011).

In CBM trainings, the emotional valence of the critical stimuli (e.g. pictures of angry faces, spiders, or alcoholic drinks) is almost always presented as a task-irrelevant feature. For instance, participants are not asked to respond to which emotion a face shows (angry or happy), which animal (spider or butterfly) is depicted in the picture, or which drink (alcoholic or soda) a picture displays. Instead, participants are instructed to respond to some physical feature of the stimuli when their approach-avoidance tendencies are trained (e.g. the grey vs. sepia colour of face pictures or the portrait vs. landscape format of beverage pictures), or there are no instructions at all, as in the dot-probe task which is widely used to train attentional processes. These trainings rely on the assumption that despite being task-irrelevant, the emotional valence of the stimuli will be processed, such that a training effect at the emotional level, rather than at the physical level alone, can be achieved. It seems that in some CBM studies, this was indeed the case because the training affected emotional vulnerability, anxiety symptoms, or drug abstinence after the training. However, it is unknown whether processing of task-irrelevant emotional information also occurred in CBM studies that failed to affect the targeted cognitive process and / or clinical outcome variables (Clarke, Notebaert, & MacLeod, 2014; Vandenbosch & De Houwer, 2011; Woud, Becker, Lange, & Rinck, 2013). Unfortunately, there is no systematic research that tried to determine if and to which degree the
emotional expression of faces is processed during CBM procedures when the task requires participants to respond to another, usually physical feature of the training stimuli instead of the expression itself.

Finally, there is reason to doubt that emotional information will generally be processed with top priority when we take another aspect of information processing into account: efficiency. According to the Multimode Model of Attention (Johnston & Heinz, 1978), attentional selection is flexible and can take place at several stages. The selection could happen at an early stage (in line with Broadbent, 1958), in which it is based on basic physical features such as colour, location, or size. It could also happen at a later stage (in accord with Treisman, 1960), where the selection would be based on meaning or content. Although the selection could happen at both early and late stages, the costs are apparently different because later selection needs more cognitive capacity. This is illustrated by, for instance, a study by Johnston and Heinz (1978), in which participants took longer to respond when verbal messages differed in meaning, than when they differed physically. Johnston and Heinz (1978) concluded that processing of the meaning requires more cognitive capacity than processing of physical surface features. A similar conclusion can be drawn from experiments on visual attention, which suggest that the processing of basic physical stimulus features (e.g. colour or location) is easier and faster than the processing of meaning or content (Driver, 2001). Taken together, these results suggest that colour rather than emotion should enjoy a processing advantage.

Therefore, the present study was designed to test these opposing predictions and to find out (1) how task-irrelevant features are processed by participants in an AAT; and (2) whether emotions are processed with higher priority than simple physical features, because of their high ecological significance. In three experiments, we presented pictures of faces that contained both emotional information (angry or happy facial expression) and basic physical information (grey vs. sepia colour of the picture). We used the approach-avoidance task mentioned above to make only one of these features task-relevant by asking participants either to respond to picture valence or to respond to picture colour. Unbeknownst to the participants, the two features were correlated, such that approaching/avoiding a specific emotion implied approaching/avoiding a specific colour, and vice versa. The question of feature priority was addressed by determining whether participants would show a training effect for task-irrelevant emotional information when physical information was task-relevant and trained, and vice versa. In Experiment 1, this was studied without additional measures that might draw participants’ attention to one or the other stimulus feature. In Experiment 2, so-called NoGo-trials were added to make sure that both the task-relevant and the task-irrelevant stimulus feature would be attended to, making it easier to process the contingency of colour and emotion. Finally, Experiment 3 directly compared a training with NoGo-trials to a training without them, using both word and picture stimuli.

**Experiment 1: processing of emotion information versus colour information**

In this AAT training, participants were trained to approach or avoid pictures of faces. The pictures always combined an emotional feature with a physical feature because the faces looked happy or angry, and they were printed in grey or brown colour. However, for each participant only one of the two features was task-relevant: Half of the participants were instructed to react to the colour (pulling one colour closer and pushing the other colour away) and the other half reacted to the emotion (pulling one emotion closer and pushing the other emotion away). During the training, the two features emotion and colour were perfectly correlated with each other, such that approaching one emotion and avoiding the other emotion also meant approaching one colour and avoiding the other colour. Whether training of the task-relevant feature affected learning of the task-irrelevant feature was measured with brief test blocks inserted after each training block. These test blocks contained colour-emotion combinations which were the opposite of what had been trained. If learning of the trained combinations occurs, reaction times to these reversed combinations should increase across training. We hypothesized that the task-irrelevant feature (either emotion or colour) could be trained, despite the fact that the participants were instructed to react to the relevant feature only. Moreover, we expected that the learning effect would be more pronounced for task-irrelevant emotion (when participants reacted to colour) than for task-irrelevant colour (when participants reacted to emotion). This prediction was based on the high relevance that emotional information had shown in previous studies (Rinck et al., 2013; Taylor & Amir, 2012).
Methods

Participants
We recruited 185 participants (mean age 21.5 years, range 18–64), mainly students of Radboud University Nijmegen. The sample was from the general student population and colour blindness was the only exclusion criteria. After completing the experiment, the participants received 1 h of course credits or 10 Euro. Participants were randomly assigned to one of four groups: In the “emotion-relevant, intuitive” group, participants were instructed to respond to emotion by approaching happy faces (which were always grey) and avoiding angry faces (which were always brown). This emotion-motion combination is in line with people’s intuitive responses (e.g. Heuer et al., 2007). In the “color-relevant, intuitive” group, participants were instructed to respond to colour by approaching grey faces and avoiding brown faces. Since the approached grey faces were always happy, and the avoided brown faces were always angry, this group responded in the same intuitive way to emotional expressions as the first group. In contrast, participants in the “emotion-relevant, counterintuitive” group and the “color-relevant, counterintuitive” group did the opposite: they approached brown-angry faces and avoided grey-happy faces, according to instructions that made emotion or colour task-relevant, respectively. The four groups did not differ significantly in size (48 colour-intuitive, 47 colour-counterintuitive, 45 emotion-intuitive, 45 emotion-counterintuitive), age, F(3,181) = 1.05, p = .37, gender, χ²(3) = 1.89, p = .60, or social anxiety, F(3,181) = 2.33, p = .08.

Materials

Approach-avoidance task. The AAT was adapted from the task used by Heuer et al. (2007). In the current task, single stimuli (pictures of emotional faces in colour) were presented on a computer screen. The participants’ task was to respond as quickly as possible to each picture by pushing or pulling a joystick attached to the table in front of the participant. As recommended, a zoom function was used: When the joystick was pulled, the picture increased in size until it filled the screen in height. When the joystick was pushed, the picture decreased in size until it became extremely small (Heuer et al., 2007). The picture disappeared only after a full movement of the joystick in the correct direction was made. The pictures used in the task depicted happy and angry faces, selected from the Radboud Faces Database (RaFD; Langner et al., 2010) and from the Karolinska Directed Emotional Faces (KDEF; Lundqvist, Flykt, & Öhman, 1998). In total, 12 male and 12 female models were selected. Each model contributed both a happy and an angry expression. The pictures from the two databases were modified to be comparable in size, brightness, tone and background. A grey-scale version and a brown (“sepia”) version were generated of each picture. The AAT consisted of eight blocks. In each block, there was a training phase and a test phase, in order to track the learning curve taking place across the whole training blocks.

Training blocks. The task consisted of 8 training blocks of 48 training trials each. In each of these blocks, participants of the two “intuitive” groups always approached grey-happy faces 24 times and avoided brown-angry faces 24 times. Participants of the two “counter-intuitive” groups did the opposite, approaching brown-angry faces 24 times and avoiding grey-happy faces 24 times.

Test blocks. Each training block was followed by a brief test block of 8 test trials. Of these, 4 trials were compatible with the training trials, showing the same colour-emotion combinations (grey-happy, brown-angry). In the remaining 4 training-incompatible trials, the combinations were reversed (grey-angry, brown-happy). This did not affect the instructions: Participants in the “emotion” groups still responded to the emotion, and participants in the “colour” groups still responded to the colour. It did, however, introduce a reversal of the task-irrelevant information. This reversal was introduced to test whether participants would learn the feature combinations presented during the training trials: If they did, a change of the task-irrelevant information should yield an increase in response times, even though the instructions encouraged participants to ignore all task-irrelevant features. In total, there were 8 of these 8-trial test blocks.

Liebowitz social anxiety scale (LSAS; Liebowitz, 1987). The LSAS measures participants’ level of social anxiety. It has high validity (Heimberg et al., 1999). In the present study, only the fear-subscale was employed. The scale contains 24 items which describe varied social situations. The participants have to indicate to which extent they are afraid of each situation, using a 4-point Likert scale (“none” to “severe”). In the present study, participants’ social anxiety level was used as a control variable, because in earlier studies, social anxiety correlated positively with avoidance of angry faces (Van Peer, Spinhoven,
van Dijk, & Roelofs, 2009), and avoidance of smiling faces (Heuer et al., 2007; Lange, Keijsers, Becker, & Rinck, 2008; Roelofs et al., 2010).

Contingency awareness check. In order to assess how much participants became aware of the experimental contingencies, they were asked four questions about the movements associated with the two emotions and the two colours: “How often did you pull or push the pictures with happy faces / angry faces / gray pictures / brown pictures?”. Participants gave their answers using visual analogue scales ranging from “100% pull” to “100% push”, with “50% pull 50% push” in the middle. For the analyses, the correct proportions were subtracted from participants’ answers, and difference scores close to zero indicate a high level of contingency awareness.

Procedure
Participants first gave informed consent to the study, then completed the LSAS on a computer. Afterwards, they were randomly assigned to one of the four experimental groups. Depending on the group, they were instructed to pull or push the joystick according to the emotional expression of the faces, or according to the colour of the pictures. There were four self-paced breaks in the AAT. The whole task including LSAS and AAT training lasted about 40 min.

Results
Data preparation
Before the analyses, two participants had to be excluded because their data were missing due to technical problems. This left 183 participants for the analyses reported below. To find out whether participants are sensitive to the task-irrelevant features during the AAT training, an irrelevant feature compatibility-score was computed for each participant and each of the 8 test blocks. The score was computed by subtracting the mean RT of the 4 training-compatible trials (grey-happy, brown-angry) from the mean RT of the 4 training-incompatible trials (grey-angry, brown-happy). A positive compatibility-score indicates that in this test block, the participant responded more quickly to the “usual” colour-emotion combinations (grey-happy, brown-angry) than to the reversed combinations (grey-angry, brown-happy), even though one of the two features, and therefore also the combination of features, was task-irrelevant and could safely be ignored. This way, positive scores indicate that the task-irrelevant feature is being processed, and increasing scores over the course of the training indicate increasing learning of the task-irrelevant colour-emotion combinations. After computing the score for each test block, we averaged the scores across consecutive block pairs (1&2, 3&4, 5&6, 7&8), yielding 4 more stable scores for the analyses, in order to reduce noise in the data (Cronbach’s α = .77). No participants’ compatibility-score was detected as an outlier (3 SDs above/below the mean), so all 183 participants were kept in the analyses.

Compatibility-scores
The irrelevant-feature compatibility-scores were analyzed using a 2 × 2 ANOVA with the between-subjects factors “Task relevance” (emotion vs. colour) and “Intuitiveness” (intuitive vs. counter-intuitive). This analysis did not yield any differences between the intuitive and the counter-intuitive conditions, \( F(1,179) = .74, p = .39, \eta^2_p = .004 \). The factor “Task relevance” was highly significant, \( F(1,179) = 35.37, p < .001, \eta^2_p = .17 \), indicating that the compatibility-scores were larger in the emotion-relevant groups than in the colour-relevant groups. The interaction of “Task relevance” and “Intuitiveness” was not significant, \( F(1,179) = 1.51, p = .22, \eta^2_p = .008 \), suggesting that the difference between emotion-relevant and colour-relevant groups did not depend on whether the movements were intuitive or counter-intuitive. The mean compatibility-scores are shown in Table 1. Finally, all analyses reported so far were repeated with the participants’ LSAS scores used as a covariate. Highly similar results were observed.

Reaction times
To explore the time course of learning effects in the different conditions in more detail, reaction times were also analyzed in addition to compatibility-scores (see Table 1). To that end, a 2 × 4 × 2 mixed-factors ANOVA was computed with the between-subjects factors “Task relevance” and the within-subjects factors “Time” and “Compatibility” (training-compatible vs. training-incompatible trials). The significant main effect of “Task relevance” revealed that the emotion-relevant groups were generally slower than the colour-relevant groups, \( F(1,181) = 120.73, p < .001, \eta^2_p = .40 \). The critical interaction of “Compatibility” and “Time” was also significant, \( F(3,543) = 12.63, p < .001, \eta^2_p = .07 \). Inspection of the mean RTs showed that RTs for the training-compatible trials kept stable across
blocks (linear contrast: $F(1,181) = .36, p = .55, \eta^2_p = .002$), whereas RTs for training-incompatible trials increased over time (linear contrast: $F(1,181) = 34.89, p < .001$, $\eta^2_p = .16$). Hence, changes in compatibility-scores stem mainly from the RT increase for training-incompatible trials. This pattern was observed similarly for the emotion-relevant groups (significant linear contrast for incompatible trials, $F(1,87) = 8.11, p = .005$, $\eta^2_p = .085$, but not for compatible trials, $F(1,87) = 2.11, p = .15$, $\eta^2_p = .02$) and for the colour-relevant groups (larger significant linear contrast for incompatible trials, $F(1,94) = 31.85, p < .001$, $\eta^2_p = .25$, than for compatible trials, $F(1,94) = 7.54, p = .007$, $\eta^2_p = .07$).

**Table 1.** Means (and standard deviations) in ms of compatible trials, incompatible trials, and the resulting compatibility-scores in each block pair of Experiment 1.

<table>
<thead>
<tr>
<th></th>
<th>Block 1&amp;2</th>
<th>Block 3&amp;4</th>
<th>Block 5&amp;6</th>
<th>Block 7&amp;8</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Compatible trials</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colour-relevant groups</td>
<td>585 (70)</td>
<td>603 (75)</td>
<td>598 (68)</td>
<td>602 (70)</td>
</tr>
<tr>
<td>Emotion-relevant groups</td>
<td>679 (87)</td>
<td>689 (85)</td>
<td>672 (74)</td>
<td>674 (83)</td>
</tr>
<tr>
<td><strong>Incompatible trials</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colour-relevant groups</td>
<td>590 (106)</td>
<td>622 (115)</td>
<td>641 (129)</td>
<td>668 (137)</td>
</tr>
<tr>
<td>Emotion-relevant groups</td>
<td>762 (137)</td>
<td>803 (113)</td>
<td>786 (124)</td>
<td>815 (139)</td>
</tr>
<tr>
<td><strong>Compatibility-scores</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colour-relevant groups</td>
<td>5 (79)</td>
<td>19 (75)</td>
<td>43 (96)</td>
<td>67 (111)</td>
</tr>
<tr>
<td>Emotion-relevant groups</td>
<td>82 (149)</td>
<td>114 (130)</td>
<td>112 (140)</td>
<td>140 (164)</td>
</tr>
</tbody>
</table>

**Discussion**

The present study tried to illuminate how participants process stimuli with emotional features and colour features in a task in which only one of these features is relevant. The task consisted of a joystick training procedure using the Approach-Avoidance Task (Heuer et al., 2007). In this task, half of the participants had to respond to the emotional expression of faces by pulling or pushing a joystick, and they were free to ignore the colour of the face pictures. The other participants had to move the joystick in response to the colour of the face pictures, and they were free to ignore the emotional expression. Nevertheless, the results of the test blocks that were inserted into the training procedure suggested that, as we had hypothesized, participants did not ignore the task-irrelevant stimulus features. However, contradictory to our second hypothesis, the emotional feature was not processed with high priority during the task. Instead, task-irrelevant colour had a stronger impact on participants who responded to emotion than task-irrelevant emotion had on participants who responded to colour. Moreover, the results of the awareness check suggest that the difference between the two Task Relevance conditions was not due to the participants’ awareness of the task contingencies. Moreover, participants were trained equally well in the “intuitive” and the “counterintuitive” conditions. Finally, it is difficult to draw conclusions about the participants’ pre-training tendencies in response to the two emotions anger and happiness, due to the lack of a pre-test. Since this question was not the focus of the current experiment, we can only assume that the participants learnt quite quickly during the training.

**Experiment 2: emotion information vs. colour information, with NoGo-trials**

Experiment 1 showed that task-irrelevant features were processed during an approach-avoidance training. However, emotional information was not

---

**Awareness check**

The participants’ awareness of the relation between the movements and the task-irrelevant features was assessed for the colour-relevant groups (awareness of emotion contingency) and the emotion-relevant groups (awareness of colour contingency) separately. Then a univariate ANOVA was employed with “Task relevance” and “Intuitiveness” as the between-subject factors and the awareness difference scores as the dependent variable. The significant main effect of “Task relevance” revealed that participants in the colour-relevant groups were more aware of how they had responded to the task-irrelevant emotions than the emotion-relevant groups were aware of how they had responded to the task-irrelevant colours, $F(1,175) = 4.08, p = .05, \eta^2_p = .02$. The factor “Intuitiveness” and the interaction were not significant, $F(1,175) = .15, p = .70, \eta^2_p = .001, F(1,175) = 2.00, p = .16, \eta^2_p = .01$, respectively. Moreover, in the colour-relevant groups, a significant correlation between the awareness difference scores for the task-irrelevant emotion and the mean compatibility-score across blocks was observed: Greater awareness (i.e. smaller differences) was linked to a greater training effect ($r(92) = -.30, p = .004$). No correlation was found in the emotion-relevant groups ($r(85) = -.04, p = .73$).
processed with high priority. This finding can be explained by the Multimode-Model of Attention (Johnston & Heinz, 1978). This model suggests that the processing costs of content features such as emotional expressions are higher than those of physical features because processing of the former consumes more resources. Hence, in order to work most efficiently, it is possible to give the more easily processed feature a higher priority. In the AAT, participants are instructed to move the joystick as quickly as possible. Under such time pressure, the task can be accomplished better by ignoring the more difficult feature than by ignoring the easier feature. In Experiment 1, this was the case when the task itself did not require processing of emotional features (in the colour-relevant conditions). However, the Multimode-Model also assumes that when participants are required to process the task-irrelevant stimulus feature, the more difficult feature (here: emotion in the colour-relevant conditions) will reach the same level of priority as the easier feature (here: colour in the emotion-relevant conditions). To test this prediction, Experiment 2 was designed. It was similar to Experiment 1, but additional so-called NoGo-trials were presented among the training trials. On these NoGo-trials, the participants were required to not move the joystick for 2 s. After that, the picture would disappear by itself. The NoGo-trials forced participants to pay attention to the task-irrelevant feature as well. Forcing all participants to process the task-irrelevant feature was expected to decrease the difference in the compatibility-scores we had observed between the emotion-relevant and colour-relevant conditions. In Experiment 2, we predicted that participants in the colour-relevant group would learn to react to emotion as quickly as participants in the emotion-relevant group learned to react to colour. Since Experiment 1 had not yielded differences between the intuitive and counter-intuitive conditions, and since the intuitive conditions have practical significance (see Taylor & Amir, 2012), Experiment 2 contained only the two intuitive conditions.

Methods

Participants

We recruited 108 participants (mean age 22.2 years, range 18–65), mainly students of Radboud University Nijmegen. The sample came from the general student population, and impaired colour recognition was the only exclusion criteria. After completing the experiment, participants received 0.5 h of course credit. Each participant was randomly assigned to one of two groups which differed in the task-relevant stimulus feature (emotion-relevant vs. colour-relevant). Hence, half of the participants were instructed to react to the colour of the faces, while the other half reacted to the facial expression. The two groups did not differ in size (n = 54 in both groups), age, t(106) = 1.52, p = .13, or gender distribution, X²(1) = .07, p = .79.

Materials

The materials were similar to those of Experiment 1. However, we did not measure participants’ social anxiety with the LSAS, because no moderating effect of social anxiety had been found in Experiment 1.

Approach-avoidance task. The task was similar to the one employed in Experiment 1. There were again eight training blocks.

Training blocks. Different from Exp. 1, each training block contained 8 NoGo-trials, randomly interspersed among the 48 training trials. In these NoGo-trials, an additional variation of the task-irrelevant feature was presented. In the colour-relevant group, the NoGo-trials showed faces with a surprised expression. In the emotion-relevant group, the NoGo-trials showed faces in blue colour. Participants were instructed to respond to these NoGo-trials by holding the joystick in the middle position for 2 s. Since NoGo-trials were randomly mixed with training trials, participants needed to constantly process the task-irrelevant feature: It determined whether the participant had to respond at all, while the task-relevant feature determined the nature of the response, pulling versus pushing.

Test blocks. As in Experiment 1, each training block was followed by a brief test block consisting of 4 trials with the untrained face-colour combination and 4 trials with the trained face-colour combination. In addition, a longer pre-test was administered before the first training block. This pre-test consisted of 24 trials with the trained face-colour combination and 24 trials with the untrained combination. These pre-test trials served to yield a baseline, pre-training compatibility-score. We expected the score to be around 0, since no training was applied yet.

Awareness check. The questions were similar to the ones in Exp. 1: Participants were asked how often they had pulled or pushed the pictures with happy faces, with angry faces, the grey pictures, and the brown pictures.
**Procedure**

After the participants arrived at the laboratory, they were randomly assigned to one of the two instruction groups and they received the corresponding instructions for the AAT. Participants in the colour-relevant group were asked to pull grey pictures closer and to push brown pictures away, but they should not react to pictures with a surprised facial expression. Participants in the emotion-relevant group were asked to pull smiling faces closer and push angry faces away, but they should not react to faces in blue colour. For both groups, responses to these NoGo-trials were counted as mistakes. The whole task lasted approximately 30 min.

**Results**

**Compatibility-scores**

The compatibility-scores were computed in the same way as in Experiment 1, and their internal reliability was again fairly high for a reaction time task (Cronbach’s α = .49). Five participants’ compatibility-scores were detected as outliers. Their scores were transformed by using the winsorization method, and then kept in the analyses. The resulting compatibility-scores are shown in Table 2. Similar to Experiment 1, the scores were analyzed using a mixed-factors ANOVA with the between-subjects factor “Task relevance” (emotion vs. colour) and the within-subjects factor “Time” (pre-test, training). For the latter, a mean compatibility score for the complete training phase was computed. The analysis revealed a significant main effect of “Time”, $F(1,106) = 110.66, p < .001$, $\eta^2_p = .51$, suggesting that the compatibility-scores were larger during the training than during the pre-test. The main effect of “Task relevance” was also significant, $F(1,106) = 5.46, p = .02$, $\eta^2_p = .05$, indicating that – as in Experiment 1 – the compatibility-scores were larger in the emotion-relevant group than in the colour-relevant group. Unlike Exp. 1, the interaction of “Time” and “Task relevance” was also significant, $F(1,106) = 11.61, p = .001$, $\eta^2_p = .10$, because the compatibility-scores increased more from pretest to training in the emotion-relevant group than in the colour-relevant group.

**Reaction times**

As in Experiment 1, reaction times were analyzed in addition to compatibility-scores, in order to assess the time course of learning in more detail (see Table 2). To that end, a $2 \times 5 \times 2$ mixed-factors ANOVA was computed with the between-subjects factors “Task relevance” and the within-subjects factors “Time” and “Compatibility” (training-compatible vs. training-incompatible trials). The significant main effect of “Task relevance” revealed that as in Exp. 1, the emotion-relevant group was generally slower than the colour-relevant group, $F(1,106) = 6.87, p = .01$, $\eta^2_p = .06$. The interaction of “Compatibility” and “Time” was also significant, $F(4,424) = 18.50, p < .001$, $\eta^2_p = .15$. Inspection of the mean RTs showed that RTs for the training-compatible and –incompatible trials were very similar before the training. After the first two blocks, however, RTs for the training-incompatible trials had increased dramatically (linear contrast: $F(1,107) = 33.00, p < .001$, $\eta^2_p = .24$), whereas RTs for the training-compatible trials did not increase significantly (linear contrast: $F(1,107) = 2.00, p = .16$, $\eta^2_p = .02$), and they both stayed relatively stable throughout the remaining blocks.

**Awareness check**

As in Experiment 1, a numerical score for the participants’ awareness of the contingency between the task-irrelevant feature and the movement directions was calculated. These scores were subjected to a univariate ANOVA with “Task relevance” as the between-subjects factor. The main effect of “Task relevance” was not significant, $F(1,106) = 2.95, p = .09$, $\eta^2_p = .03$, suggesting that the two groups had similar levels of contingency awareness. Moreover, a significant correlation between contingency awareness and

| Table 2. Means (and standard deviations) in ms of compatible trials, incompatible trials, and resulting compatibility-scores in pre-test and training block pairs of Experiment 2. |
|--------------------|----------------|----------------|----------------|----------------|----------------|
|                    | Pre-Test       | Block 1&2      | Block 3&4      | Block 5&6      | Block 7&8      |
| Compatible trials  |                |                |                |                |                |
| Colour-relevant group | 540 (70)     | 652 (88)       | 640 (85)       | 616 (77)       | 616 (81)       |
| Emotion-relevant group | 661 (118)    | 644 (90)       | 644 (87)       | 639 (86)       | 632 (82)       |
| Incompatible trials |                |                |                |                |                |
| Colour-relevant group | 547 (78)     | 700 (130)      | 671 (122)      | 690 (112)      | 678 (161)      |
| Emotion-relevant group | 656 (119)    | 744 (142)      | 733 (125)      | 733 (132)      | 711 (129)      |
| Compatibility-scores |                |                |                |                |                |
| Colour-relevant group | 6 (30)        | 47 (90)        | 32 (67)        | 74 (77)        | 60 (105)       |
| Emotion-relevant group | −3 (35)      | 98 (97)        | 85 (90)        | 94 (86)        | 79 (89)        |
compatibility scores was found in the emotion-relevant group, suggesting that greater awareness was linked to a larger training effect, \( r(54) = -0.47, p < .001 \). No correlation was found in the colour-relevant group, \( r(54) = 0.07, p = .62 \).

**Discussion**

In Experiment 2, NoGo-trials were introduced into the training phase, forcing participants to process the task-irrelevant feature as well. This way, emotional information became relevant for the decision whether to react at all, while it remained irrelevant for the type of response, that is, approach or avoidance. The addition of the NoGo-trials was meant to decrease, and possibly eradicate, the observed difference between the effects for task-irrelevant colour and task-irrelevant emotion. Indeed, the difference between the two conditions was greatly reduced in Experiment 2. Learning of task-irrelevant emotional information seemed to take longer than learning of task-irrelevant colour information, but there was no significant difference anymore in the second half of the training. Thus, Experiment 2 suggests that emotional information can be learned just like colour information, but both Experiment 1 and 2 suggest that emotion information is not processed as automatically as assumed previously.

**Experiment 3: emotional pictures and words, with versus without NoGo-trials**

When comparing the results of Experiment 2 to those of Experiment 1, it seems that the learning of task-irrelevant emotional information can indeed be boosted by the addition of NoGo-trials which force participants to process emotional information. On average, the compatibility effect for task-irrelevant emotion information increased from 33 ms in Exp. 1 to 58 ms in Exp. 2. However, at the same time, the compatibility effect for task-irrelevant colour information decreased from 112 ms to 93 ms, possibly explaining why in the second half of the training, the effects of colour and emotion were not significantly different anymore. This illustrates how difficult it is to draw conclusions from a comparison of two different experiments run separately at different times. Therefore, we decided to conduct Experiment 3, in which we systematically compared the learning of task-irrelevant emotional information in a training with NoGo-trials (as in Exp. 2) to a training without them (as in Exp. 1). In addition, Experiment 3 also addressed how stimulus-specific the learning effect was by investigating not only emotional pictures (the happy and angry faces used before) but also emotional words (positive vs. negative). In all cases, participants were instructed to respond with approach or avoidance to stimulus colour (again: grey or brown), while emotional valence was irrelevant for the response type. We predicted that for both pictures and words, participants in the “with NoGo-trials” condition would learn the emotional feature better than participants in the “without NoGo-trials” condition.

**Methods**

**Participants**

We recruited 206 participants (mean age 22.7 years, range 17–68) from the general student population. Most participants were students of Radboud University Nijmegen, and impaired colour recognition was the only exclusion criteria. After completing the experiment, the participants received 0.5 h of course credit. They were randomly assigned to one of four conditions. Hence, half of the participants reacted to the colour of emotional pictures (faces), while the other half reacted to the colour of emotional words. Within each group, half of the participants had NoGo-trials during their training, while the other half did not. The resulting four groups did not differ in size (always 51 or 52 participants), in age, \( F(3,205) = .60, p = .60 \), or in gender distribution, \( \chi^2(3) = 3.42, p = .33 \).

**Materials**

**Approach-avoidance task.** The “picture-with-NoGo” condition was identical to the one used in Experiment 2. The “picture-without-NoGo” condition was the same, except that it did not contain any NoGo-trials. For the two conditions involving word stimuli, the structure of the training was the same, except that the face pictures were replaced by emotional words. These words were either positive or negatively valenced Dutch words. All words were nouns, adjectives, or verbs, containing 5–8 letters (e.g. kill, violence, friend, healthy). In total, there were 28 positive words and 28 negative words. Like the pictures, the words were printed in grey or brown colour, and the zoom function was used for them as well. As in Experiment 2, the training consisted of a pre-test block followed by 8 training blocks. The whole task lasted approximately 30 min.
**Awareness check.** The questions were similar to the ones in Experiments 1 & 2; participants were asked how often they had pulled or pushed the pictures with happy faces, with angry faces, the grey pictures, and the brown pictures. However, different from the other two experiments, colour was the task-relevant feature in all four conditions. Hence, the awareness check also served as a manipulation check, to assess whether the two conditions with NoGo trials would show better awareness of the contingency between motion direction and the task-irrelevant feature “emotion”.

**Results**

**Awareness check**

The participants’ contingency awareness scores regarding the emotion feature were calculated as in Experiments 1 and 2. One participant’s compatibility-score was detected as an outlier, hence this score was transformed by using winsorization and kept in the analyses. Then A univariate ANOVA was employed with “Stimuli (words vs. pictures)” and “Presence of NoGo trials (with vs. without)” as the between-subjects factors. The results showed that participants in the “with NoGo-trials” conditions had a higher contingency awareness than those in the “without NoGo-trials” conditions, $F(1,199) = 436.40, p < .001, \eta^2_p = .69$, participants in “words” conditions had a higher awareness than those in “pictures” conditions, $F(1,199) = 9.63, p = .002, \eta^2_p = .05$. The interaction between “Stimuli” and “Presence of NoGo trials” was also significant, $F(1,199) = 13.84, p < .001, \eta^2_p = .07$. The significant interaction revealed that when NoGo trials were added to the training, participants reacting to words had a better awareness of the emotion continuity than participants reacting to picture stimuli, while without NoGo trials, the stimuli did not make a difference for the participants’ awareness. Furthermore, a significant correlation between awareness scores and compatibility-scores was only found in the “pictures-with-NoGo” group, $r(52) = -.35, p = .01$: Higher awareness went along with a greater training effect. In the other three conditions, the correlations remained insignificant, $|r| < .14, p > .33$.

**Compatibility-scores**

The compatibility-scores shown in Table 3 were computed in the same way as in the previous experiments. The internal reliability of the compatibility-scores was very high for a reaction time task (Cronbach’s $\alpha = .71$).

We first computed a mixed-factors $2 \times 2 \times 2$ ANOVA of the compatibility-scores with the between-subjects factors “Stimuli” and “Presence of NoGo-trials”, and the within-subjects factor “Time” (pre-test, training). The significant main effect of “Time”, $F(1,202) = 196.89, p < .001, \eta^2_p = .60$, showed that the compatibility-scores increased from the pre-test to the training. The main effect of “Presence of NoGo-trials” was significant as well, $F(1,201) = 77.15, p < .001, \eta^2_p = .28$: In the presence of NoGo trials, the compatibility-scores were larger than in their absence. This effect of the NoGo-trials was larger for words than for pictures, as indicated by a significant two-way interaction of “Stimuli” and “Presence of NoGo-trials”, $F(1,202) = 17.01, p < .001, \eta^2_p = .08$. Finally, the interaction of all three factors was also significant, $F(1,202) = 16.36, p < .001, \eta^2_p = .08$, suggesting that the enhancing effect of NoGo-trials on learning was stronger for words than for pictures (see Table 3).

To better understand the effects observed for pictures and words, the two stimulus types were analyzed separately. For pictures, both the main effect of time and the main effect of “Presence of NoGo-

| Table 3. Means and standard deviations in ms of compatible trials, incompatible trials, and resulting compatibility-scores in pre-test and training block pairs of Experiment 3. |
|---------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                                 | Pre-Test | Block 1&2 | Block 3&4 | Block 5&6 | Block 7&8 |
| **Compatible trials**           |          |          |          |          |          |
| Words with NoGo-trials          | 531 (64) | 708 (83) | 662 (69) | 649 (71) | 643 (73) |
| Words without NoGo-trials       | 546 (68) | 557 (63) | 548 (62) | 541 (58) | 541 (57) |
| Pictures with NoGo-trials       | 542 (71) | 634 (85) | 617 (86) | 598 (77) | 596 (76) |
| Pictures without NoGo-trials    | 557 (73) | 567 (71) | 565 (66) | 562 (81) | 563 (84) |
| **Incompatible trials**         |          |          |          |          |          |
| Words with NoGo-trials          | 527 (62) | 883 (202)| 812 (174)| 828 (182)| 813 (169)|
| Words without NoGo-trials       | 541 (65) | 589 (98) | 575 (90) | 568 (80) | 606 (114)|
| Pictures with NoGo-trials       | 543 (71) | 718 (142)| 692 (135)| 719 (138)| 690 (118)|
| Pictures without NoGo-trials    | 557 (83) | 602 (118)| 599 (104)| 623 (141)| 622 (121)|
| **Compatibility-scores**        |          |          |          |          |          |
| Words with NoGo-trials          | -4 (20)  | 176 (155)| 150 (144)| 179 (140)| 169 (144)|
| Words without NoGo-trials       | -5 (18)  | 33 (64)  | 27 (73)  | 27 (62)  | 65 (93)  |
| Pictures with NoGo-trials       | 1 (24)   | 83 (93)  | 75 (106) | 121 (114)| 94 (94)  |
| Pictures without NoGo-trials    | 0 (21)   | 35 (86)  | 34 (86)  | 60 (110) | 59 (90)  |
trials” were significant, $F(1,101) = 109.89, \ p < .001, \ \eta^2_p = .52$; and $F(1,101) = 14.01, \ p < .001, \ \eta^2_p = .12$, indicating that the compatibility-scores were larger during training than during pre-test, and that the presence of NoGo-trials increased the compatibility-scores (see Table 3). Moreover, the interaction of these two factors was significant, $F(1,101) = 11.69, \ p = .001, \ \eta^2_p = .10$, indicating that the presence of NoGo-trials boosted the learning of task-irrelevant picture features.

For words, the main effects of “Time” and of “Presence of NoGo-trials” as well as their interaction were significant, $F(1,101) = 187.81, \ p < .001, \ \eta^2_p = .65$; $F(1,101) = 67.99, \ p < .001, \ \eta^2_p = .40$; and $F(1,101) = 67.58, \ p < .001, \ \eta^2_p = .40$, respectively. The pattern of means shown in Table 3 suggests that the compatibility-scores increased from pre-test to training, that NoGo-trials increased the compatibility-scores, and that in the presence of these trials, the increase of the scores from pre-test to training was larger than in their absence.

**Reaction times**

As in Experiments 1 and 2, reaction times were analyzed in addition to compatibility-scores, in order to assess the time course of learning effects. A $2 \times 2 \times 5 \times 2$ mixed-factors ANOVA was computed with the between-subjects factors “Stimuli” and “Presence of NoGo-trials” and the within-subjects factors “Time” and “Compatibility”. The main effects of “Presence of NoGo-trials” and “Stimuli” were both significant: Participants reacted more slowly in the presence of NoGo-trials, $F(1,202) = 85.07, \ p < .001, \ \eta^2_p = .30$, and more slowly to words than to pictures, $F(1,202) = 5.53, \ p = .02, \ \eta^2_p = .03$. As in the previous experiments, a significant interaction of “Compatibility” and “Time” was also observed, $F(3,606) = 3.53, \ p = .015, \ \eta^2_p = .02$. Linear contrasts revealed the nature of this interaction: The RTs for training-compatible trials showed significant linear decreasing trends in almost all training conditions (pictures-with-NoGo: $F(1,51) = 51.84, \ p < .001, \ \eta^2_p = .50$; words-with-NoGo: $F(1,51) = 125.55, \ p < .001, \ \eta^2_p = .71$; words-without-NoGo: $F(1,50) = 11.63, \ p = .001, \ \eta^2_p = .19$, pictures-without-NoGo, $F(1,50) = .45, \ p = .51, \ \eta^2_p = .01$). In contrast, for the training-incompatible trials, nearly all training conditions did not show any trends over time (pictures-with-NoGo: $F(1,51) = 1.40, \ p = .24, \ \eta^2_p = .03$; words-without-NoGo: $F(1,50) = .35, \ p = .35, \ \eta^2_p = .02$; picture-without-NoGo: $F(1,50) = 2.85, \ p = .10, \ \eta^2_p = .05$, words-without-NoGo, $F(1,51) = 4.46, \ p = .04, \ \eta^2_p = .08$.

**Discussion**

In Experiment 3, the beneficial effect of NoGo-trials on the learning of task-irrelevant emotional information was quite obvious. When participants had to attend to the emotional features of words and faces because these features determined whether they had to respond at all, the emotional information affected the speed of responding more strongly than in the absence of NoGo-trials. After only two blocks of training, reversed emotion-colour combinations led to slower responses even though only the colour of the stimuli was relevant for the task of pulling versus pushing. This effect was clearly learned because no such effect was observed during the pre-test before training, and it was more pronounced for words than for faces. In sum, the results of Experiment 3 supply more evidence that the learning of emotion-colour combinations can be improved by forcing participants to process the emotional information. Moreover, introducing NoGo-trials also increased participants’ awareness of the contingency between movement and the task-irrelevant emotion feature, as expected.

**General discussion**

Our aims in the current experiments were to examine: (1) how bottom-up and top-down processes are integrated by participants in the AAT, namely how participants process emotional and non-emotional features under different task instructions; and (2) whether emotions are processed with a higher priority because of their high ecological significance, compared to the simple physical feature of colour; and (3) whether the processing of emotional expressions can be enhanced by a top-down approach. Hence, in all three studies, an AAT training with stimuli that combined an emotional feature (facial emotion) with a physical feature (colour) was given to the participants. In the stimuli, the two features were correlated with each other, such that approaching a task-relevant feature dimension (e.g. smiling faces or grey pictures) necessarily involved approaching the corresponding dimension of the other, task-irrelevant feature (grey pictures or smiling faces, respectively). Training effects were computed from specific test items to assess whether participants would learn the emotion-colour contingency.

To answer the questions mentioned above, training effects for the emotional feature and the physical
feature were analyzed and compared to each other. Experiment 1 showed that participants did process the task-irrelevant features and learned the emotion-colour contingency, both when emotion was task-irrelevant and when colour was. However, opposite to our expectations, stronger evidence of learning was observed for task-irrelevant colour than for task-irrelevant emotion. In Experiment 2, NoGo-trials were added to ensure awareness of the task-irrelevant features, and the difference in learning between the two features was decreased. Moreover, the increased training effect for task-irrelevant emotion was confirmed in Experiment 3, and it was found for emotional words as well as for pictures.

The current study is one of the few that focuses on how emotional features and physical features are processed in a cognitive bias modification (CBM) paradigm, here an AAT training. In CBM studies, the majority of trainings employed emotional features (e.g. valenced pictures or positive vs. negative facial emotions) as a task-irrelevant feature. It has been assumed that emotional information can be processed automatically and effortlessly because of its high salience, giving emotional features priority over physical features such as colour, shape or orientation. This assumption has not been tested systematically yet, but some AAT training studies showed evidence to support it (e.g. Eberl et al., 2013; Wiers et al., 2011). In these trainings, participants responded to physical features of the stimuli (e.g. picture shape or colour), and paying attention to the task-irrelevant emotional feature did not help in performing the task. Hence, it seems tempting to conclude that emotional features can be processed automatically and have a high priority over physical features. However, we cannot neglect a large number of studies that failed to find effects of CBM trainings on the learning of task-irrelevant stimulus features. These failures were most frequently reported in attention bias modification studies (Clarke et al., 2014), but have also been observed in approach-avoidance trainings (Vandenbosch & De Houwer, 2011; Woud et al., 2013). Therefore, it is necessary to systematically examine how emotional information is processed in these trainings, and to which extent the processing of task-irrelevant emotional information can be promoted.

First, we found in all three experiments that task-irrelevant features could be processed in an AAT training, regardless of which feature was task-irrelevant and which one task-relevant. In the current experiments, the observed compatibility-scores were positive and kept increasing throughout the training blocks, suggesting increasing learning of the emotion-colour contingency. This result is in line with a series of contingency learning studies (Schmidt & De Houwer, 2012, 2016a, 2016b; Schmidt, Crump, Cheesman, & Besner, 2007), which revealed implicit contingency learning of colour-word combinations, even without facilitation from awareness. Our current results also help to understand the positive findings of previous studies (e.g. Eberl et al., 2013; Wiers et al., 2011) in which the AAT training affected the trained approach-avoidance tendency itself, as well as causing changes at the behavioural level. Combining all these results above, we may conclude that task-irrelevant emotional features can indeed be processed, even when this is of no obvious use for the task at hand. Given that cognitive bias modification is considered a promising new treatment of emotional dysfunctions, the implications of these findings are practically and conceptually significant.

Second, we did not find any evidence for our hypothesis that emotional facial expressions enjoy a higher priority than physical features during information processing. In Experiment 1, the learning of task-irrelevant colour in the emotion-relevant conditions was much better than the learning of task-irrelevant emotion in the colour-relevant conditions. This result contradicts both our expectations and the information-processing models introduced by Treisman (1960) and Norman (1968). These models suggest that information with high pertinence lowers the threshold for attention and can be processed with high priority. Because attending to and correctly processing emotional information are adaptive behaviours that increase an individual’s probability of survival (Lang, Bradley, & Cuthbert, 1997; Öhman, Flykt, & Esteves, 2001), emotional features should be processed with high priority. However, this was not found in the current experiments. The results of our awareness checks do not provide any evidence for the assumption that the differences between the emotion-relevant and the colour-relevant conditions were due to differences in participants’ awareness. Instead, we assume that the differences might be due to a number of reasons. For instance, processing emotional information might lead to unpleasant feelings and slow down the reactions, which might motivate participants to shift attention away from emotional information,
towards simple physical information. Second, reactions in the emotion-relevant conditions were generally slower than reactions in the colour-relevant conditions. Therefore, task-irrelevant colour might have had more time and opportunity to be processed during responding to emotions than vice versa. Third, in our experiments, what could be learned was the contingency of the task-irrelevant feature with both the task-relevant one and with the corresponding response. The fact that this particular type of learning was better for colour than for emotion does not rule out that attention was paid to both features. Experiments that address other aspects of attention, for instance, the fast detection of emotion vs. colour information might yield very different results.

On the other hand, our findings may also be related to differences in the efficiency of information processing. According to the Multi-Model of Attention (Johnston & Heinz, 1978), attention selection takes place at several stages. The selection at an early stage is assumed to be based on physical features such as colour, shape and location, while the selection at a later stage is based on meaning and content. The selection can happen at both stages, but the later selection based on meaning consumes more cognitive capacity, and therefore it may be less efficient. The current trainings, like many other CBM trainings, instructed participants to respond as quickly as possible. Hence, to achieve more efficient performance, the selection should happen as early as possible, which favours the early-stage processing of physical features over the later-stage processing of emotional features. By analyzing the mean RTs in Experiments 1 and 2, we found that responding to emotion may be more difficult than responding to colour, yielding longer RTs in the emotion-relevant groups than in the colour-relevant groups. This finding could be another piece of evidence for the efficiency theory.

Finally, Experiments 2 and 3 consistently showed that the training effect for emotion as a task-irrelevant feature can be improved dramatically by increasing awareness of the task-irrelevant feature. By adding NoGo-trials to the training procedure, the emotional information remained task-irrelevant for the type of response, that is, approach or avoidance, but it became task-relevant for the decision whether to react at all. Although this manipulation did not require or encourage participants to learn the emotion-colour contingency, it did force them to pay attention to the emotional information. According to the Multi-Model of Attention, the selection of emotion information at a later stage was forced to take place as a result of including the NoGo-trials. As a consequence, the participants also learned the emotion-colour contingency better than in Experiment 1. Interestingly, despite various differences between pictorial stimuli and word stimuli, the beneficial effect of increasing emotion awareness was found for both of them. This may indicate that the processing of emotional information shows regularities across different types of stimuli. Our finding may also have very useful practical consequences: Future studies should explore if the addition of NoGo-trials to approach-avoidance trainings and other forms of CBM can be used to improve learning and training effects, and thereby hopefully also the beneficial clinical effects of such trainings.

Several limitations of the current experiments deserve to be mentioned. First, the current experiments only addressed the processing of task-relevant and task-irrelevant stimulus features in an approach-avoidance training. Thus, nothing can be said about other trainings, for instance attention bias modifications (Clarke et al., 2014). Moreover, as mentioned above, the current experiments addressed the learning of a contingency in the stimulus materials, and we do not know whether this generalizes to the learning of other features, given the setup of the current study. On top of that, the current experiments were not designed to establish whether participants learned that the task-irrelevant feature (e.g. emotion) was correlated with the task-relevant feature (e.g. colour), or with the required response (e.g. approach-avoidance), or with both. In contrast, previous studies provide ample evidence for learning of the stimulus-response link during contingency learning (e.g. Schmidt & De Houwer, 2012).

Another limitation is that the training effect might not be maximized in the current paradigm. According to Van Dessel, De Houwer, and Gast (2016), a 100% contingency (without training-incompatible trials in the training blocks), should increase the learning effect. Moreover, Van Dessel et al. (2016) tried to prevent participants from focusing on a specific point on the screen (instead of processing the content of the pictures) by varying the size of the pictures during training trials. These adjustments might increase the training effect and could be adapted in future AAT training studies.

The third limitation is that the stimuli always combined an emotional feature with a colour feature.
Therefore, there was no comparison between emotional and non-emotional features in a situation when they were both task-irrelevant. As a result, nothing can be said about the learning of emotional information or colour information in isolation. Moreover, we should not jump to the conclusion that all emotional features are processed with a lower priority than all physical features because of their higher complexity. To understand more about the processing of the emotional features, different types of emotional features need to be examined. For instance, visually simpler emotional stimuli may be processed more easily than the emotional faces used here. Finally, we need to be careful when it comes to practical implications of the current results. We only studied the learning of emotional and physical features at the level of the approach-avoidance tendencies themselves, but we did not address any effects that such a training may have on related behaviour or emotional vulnerability. For instance, it is conceivable that, as suggested above, the inclusion of NoGo-trials will improve the learning of task-irrelevant emotional information. If this improved learning will also lead to improved effects on emotional and behavioural responses needs to be tested by future studies. For instance, Rinck et al. (2013) reduced emotional vulnerability in socially anxious participants by training them to approach smiling faces, using a task in which the emotional expression of the faces was task-irrelevant. It would be interesting to find out whether such a training would reduce emotional vulnerability even more when emotion-related NoGo-trials are included.

In summary, the current experiments are among the first to investigate in detail how emotional facial expressions are processed in AAT trainings, in the context of an interaction between top-down and bottom-up processes. Our study provides evidence that the processing of emotional features during such a training does not have higher priority than the processing of physical features. In addition, the current study also suggests that awareness of the emotional feature could promote its processing during the AAT training, which may have useful practical implications for AAT trainings and other CBM paradigms. With the help of more fundamental research, we may reach a deeper understanding of the processes involved in CBM trainings, and we may learn how to maximize the beneficial effects of CBM trainings on cognitive biases and maladaptive behaviours alike.

**Notes**

1. In order to simplify the design, we did not counterbalance colour (grey, sepia) with response (approach-avoid). This decision was based on several previous studies which had shown that there was no difference in intuitiveness or response speed for “pull-gray, push-sepia” versus “pull-sepia, push-gray”.
2. Because the compatibility-scores were calculated by subtracting RTs for compatible/trained trials from RTs for incompatible/untrained trials, there is redundant information in the ANOVAs of the compatibility-scores and the RTs. Therefore, to keep the manuscript concise, we only report RT results that provide additional insights into the data. The complete results of all analyses are available on OSF: DOI:10.17605/OSF.IO/2SHCK.

**Acknowledgments**

We are grateful to the students who participated in this study. We would also like to thank the members of the “Experimental Psychopathology and Treatment” group and the reviewers of an earlier version for helpful feedback. The Behavioural Science Institute of Radboud University Nijmegen supported the study financially.

**Disclosure statement**

No potential conflict of interest was reported by the authors.

**References**


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.

