

Addressing perceptual insensitivity to facial affect in violent offenders: first evidence for the efficacy of a novel implicit training approach

M. Schönenberg*, S. Christian, A.-K. Gaußer, S. V. Mayer, M. Hautzinger and A. Jusyte

Department of Clinical Psychology and Psychotherapy, University of Tübingen, Germany

Background. Although impaired recognition of affective facial expressions has been conclusively linked to antisocial behavior, little is known about the modifiability of this deficit. This study investigated whether and under which circumstances the proposed perceptual insensitivity can be addressed with a brief implicit training approach.

Method. Facial affect recognition was assessed with an animated morph task, in which the participants (44 male incarcerated violent offenders and 43 matched controls) identified the onset of emotional expressions in animated morph clips that gradually changed from neutral to one of the six basic emotions. Half of the offenders were then implicitly trained to direct attention to salient face regions (attention training, AT) using a modified dot-probe task. The other half underwent the same protocol but the intensity level of the presented expressions was additionally manipulated over the course of training sessions (sensitivity to emotional expressions training, SEE training). Subsequently, participants were reassessed with the animated morph task.

Results. Facial affect recognition was significantly impaired in violent offenders as compared with controls. Further, our results indicate that only the SEE training group exhibited a pronounced improvement in emotion recognition.

Conclusions. We demonstrated for the first time that perceptual insensitivity to facial affect can be addressed by an implicit training that directs attention to salient regions of a face and gradually decreases the intensity of the emotional expression. Future studies should focus on the potential of this intervention to effectively increase empathy and inhibit violent behavior in antisocial individuals.

Received 20 February 2013; Revised 22 May 2013; Accepted 23 May 2013; First published online 1 July 2013

Key words: Aggression, facial affect recognition, psychopathy, sensitivity to emotional expressions training.

Introduction

In social interactions individuals are confronted with rapidly changing facial expressions and use this information to continuously monitor intentions and emotional reactions of their interaction partners. The ability to process quickly and respond adequately to these non-verbal affective cues is vital for normal social functioning as well as for the development and maintenance of stable interpersonal affiliations. Numerous studies have suggested that deficits in the recognition or appropriate interpretation of emotional social information lie at the root of disorders such as social phobia, schizophrenia and autism (Harms *et al.* 2010; Kohler *et al.* 2010; Staugaard, 2010).

Impaired identification of affective facial cues has also frequently been linked to aggressive behavior in

antisocial populations. However, reported findings vary with regard to the specificity of the impairment, with some showing pronounced deficits in the emotion categorization of disgusted (Kosson *et al.* 2002; Sato *et al.* 2009), sad (Dolan & Fullam, 2006; Eisenbarth *et al.* 2008; Hastings *et al.* 2008), angry (Fairchild *et al.* 2010; Schönenberg *et al.* 2013) and fearful facial expressions (Blair *et al.* 2004; Montagne *et al.* 2005). The inconclusiveness may arise from methodological issues (e.g. presentation of full-blown emotional expressions *versus* morphed images with varying grades of emotional intensity, short *versus ad libitum* stimulus presentation) as well as from heterogeneity of the samples that were studied (e.g. community samples with high *versus* low trait aggression, incarcerated violent offenders, adolescents with conduct disorder, psychopaths). Nonetheless, according to a recent meta-analysis, antisocial traits and behavior are most consistently associated with deficits in recognizing fearful facial affect across studies (Marsh & Blair, 2008) and there is accumulating evidence that this impairment can be traced back to dysfunctions in neural substrates

* Address for correspondence: Michael Schönenberg, Ph.D., University of Tübingen, Department of Clinical Psychology and Psychotherapy, Schleichstraße 4, D-72076 Tübingen, Germany. (Email: michael.schoenenberg@uni-tuebingen.de)

engaged in the processing of fearful expressions, i.e. the amygdala (Raine *et al.* 2010).

With reference to the prominent violence inhibition mechanism model (Blair, 2001), facial expressions of distress are proposed to elicit empathy and inhibit aggressive behavior in healthy individuals (Marsh & Blair, 2008). Antisocial individuals, particularly those with psychopathic tendencies (e.g. lack of guilt, lack of empathy, callous use of others for one's own gain), might be insensitive to these social stop signals, a deficit most probably related to hypofunction of the amygdala (Jones *et al.* 2009; Sebastian *et al.* 2012).

Given the large amount of studies convincingly linking impairments in facial affect recognition to aggressive behavior, it seems somewhat surprising that virtually nothing is known about the modifiability of this deficit. Notably, one study demonstrated that in children with psychopathic tendencies, deficits in the ability to recognize fear might be due to attentional problems, i.e. the failure to attend to the most salient aspects of facial affect (Dadds *et al.* 2006). It has been demonstrated that this visual neglect ('fear blindness') can be temporarily reversed by explicitly directing the attentional focus to the eye region of the presented fearful faces, a technique that has also been shown to successfully correct recognition deficits in patients with amygdala damage (Adolphs *et al.* 2005).

However, correct identification of prototypical emotional expressions may not accurately reflect the actual emotion recognition performance. In everyday social interactions, facial expressions change rapidly and it is important to adapt behavior continuously to subtle signs of social cues rather than to correctly identify static full-blown expressions. Thus, we believe that it may be helpful to investigate emotion recognition performance by employing methods allowing for a fine-grained assessment of facial emotion recognition impairments (Joormann & Gotlib, 2006). To date, no studies have examined whether the attentional shift to regions of the face that carry relevant affective meaning can improve the actual perceptual sensitivity toward emotional expressions, as reflected in the correct recognition of affective displays at lower intensity levels.

Thus, the goals of the present study were twofold. First, we examined the degree of facial affect recognition impairment in antisocial violent offenders with psychopathic personality traits. In order to assess perceptual sensitivity to emotional expressions, we adopted a morphing paradigm in which the participant determines the exact onset of an emotional expression in a series of computerized movies, depicting facial expressions that slowly change from neutral to full-blown emotions. We predicted a delayed recognition of primarily fearful facial expressions in

aggressive individuals as compared with matched control subjects.

The second aim of this study was to investigate whether the proposed perceptual insensitivity to fearful facial cues could be addressed with a brief implicit training targeted at this deficit. The idea that disruptions of basal stages of social information processing can be addressed by an implicit training has been put to an extensive test in anxiety disorders, which are associated with a pronounced tendency to attend to threat-related stimuli, but not with impairments in the recognition of facial affect. Attentional bias modification (ABM) training utilizes variants of the dot-probe task (MacLeod *et al.* 1986), where brief bilateral presentations of an emotional and a neutral face are followed by a target probe in the location of one of the previously presented stimuli (Bar-Haim, 2010). Participants are to discriminate as quickly as possible between two variants of the probe (e.g. a left/right pointing arrow) without compromising accuracy. In ABM variants of the dot-probe task, target location is systematically manipulated to increase the proportion of targets appearing at the location of the intended training bias (e.g. in anxiety disorders the training protocol requires participants to shift their attention away from the threatening toward the neutral face).

We adapted a similar training protocol that was modified to address the perceptual deficits evident in aggressive individuals. We aimed to alter the threshold for the recognition of the fearful expression onset by either directing attention to an emotional expression *per se* or to significant elements of a developing emotion. To address the role of attention, we employed an ABM task, in which fearful and neutral facial expressions were presented simultaneously and the fearful face was always replaced by a target. The latter appeared in the eye region of the preceding facial stimulus, allowing us to implicitly direct participants' attention to fearful cues. To investigate whether directing attention to salient regions of the emotional face alone is sufficient to improve recognition, half of the participants completed four training sessions with full-blown fearful expressions. The other half of the participants underwent the same protocol, but the intensity level of the fearful expressions was successively decreased over the course of the four sessions; this training condition was tailored to enhance perceptual sensitivity to affective faces. After completion of the training sessions, all participants were reassessed with the morphing task to determine changes in the recognition threshold for affective facial expressions. In order to quantify possible task repetition effects on performance as measured with the animated morph task, a subsample of control participants was also tested twice. Our work group previously

demonstrated that aggressive individuals exhibit recognition deficits evident in a decreased sensitivity to facial affect (Schönenberg *et al.* 2013). Therefore, we hypothesized that mere attention allocation training toward full-blown affective expressions would be unlikely to have an effect on the emotion onset assessed with the animated morph task. We expected slightly decreased thresholds for the recognition of fearful cues across all groups, but predicted significantly improved abilities only in those individuals who have learned to shift attention to salient affective facial cues in the condition where affective intensity was manipulated.

Method

Participants

Participants were antisocial violent offenders (AVOs) who were recruited from a German correctional facility (Justizvollzugsanstalt Heimsheim) via announcements within the facility. Exclusion criteria were charges with drug-related crime, domestic violence or sexual assault as well as insufficient knowledge of the German language. A total of 45 interested AVO participants were contacted by the facility's psychological service, and experimental as well as clinical assessments were conducted in designated rooms of the facility by our research group members. One AVO participant was excluded from participation at baseline due to insufficient language skills. The final sample consisted of 44 AVOs. None of these subjects had a history of schizophrenia or suffered from mental retardation.

We were able to recruit 43 educationally and age-matched healthy controls (CTLs) from the institute's participant database. The assessment was carried out in the laboratory of the department of psychology by members of our research group. In both groups, all designated individuals participated in the study. All participants provided written informed consent and received monetary compensation for participation. The study was approved by the local ethics committee and was conducted in accordance with the Declaration of Helsinki.

Diagnostic measures

Aggressive behavior was assessed with the German version (Herzberg, 2003) of the 29-item Buss-Perry Aggression Questionnaire (BPAQ; Buss & Perry, 1992), which assesses four components of aggression: physical and verbal aggression, anger, and hostility. A German version (Eisenbarth & Alpers, 2008) of the Psychopathic Personality Inventory – Revised (PPI-R; Lilienfeld & Widows, 2005) was employed in the AVO group to assess self-reported psychopathic traits.

The instrument consists of 154 items and assesses nine different facets of the psychopathy construct in its subscales (machiavellian egocentricity, social potency, fearlessness, cold-heartedness, rebellious non-conformity, blame externalization, carefree non-planfulness and stress immunity). The PPI-R also contains a validity scale (invalid answering), which indicates insincere responding.

Stimulus material

Animated morph task

Digitized color photographs of three male models (23, 25, 71) depicting six affective states (angry, happy, fearful, sad, surprised, disgusted) as well as neutral expressions were selected from the Radboud Faces Database (Langner *et al.* 2010) based on the accuracy of emotional expressions (percentage hit rate: mean = 91.67%, *s.d.* = 9.28%). The pictures were cropped to a standard size (421 × 500 pixels) and adjusted for color and luminance with Adobe Photoshop CS4 (Adobe Systems Inc., USA). Each emotional expression of every model identity was then parametrically varied using a morphing procedure (FantaMorph software; Abrosoft, China) by blending the neutral and the affective expressions, which produced a set of 51 intensity levels (2% increment steps) ranging from 0% (neutral) to 100% (angry, fearful, happy, sad, surprised, disgusted) for each model (Fig. 1 *a*). Thus, the stimulus material for the task consisted of a total of 18 distinct sequences (3 model identities × 6 affective states). Neutral, disgusted and surprised expressions of one additional model identity (33) were used to create two sequences for the exercise trials in the same manner.

Computerized training

A total of 30 male model identities (03, 09, 10, 20, 21, 24, 28, 29, 35, 36, 38, 45–55, 59, 60, 67–70, 72, 73) depicting neutral and fearful expressions (percentage hit rate: mean = 78.70%, *s.d.* = 14.57%) were selected from the Radboud Faces Database. The pictures were cropped (400 × 517 pixels), adjusted for color and luminance (Adobe Photoshop CS4) and subsequently morphed (FantaMorph software) to obtain the stimulus material depicting five distinct intensities, i.e. 0% (neutral), 30%, 45%, 60% and 75% fearful.

Procedure

The study design included an initial baseline assessment, the administration of four training sessions and one post-training assessment. All AVOs underwent the baseline, the training and the post-training assessment. All CTLs completed the initial baseline

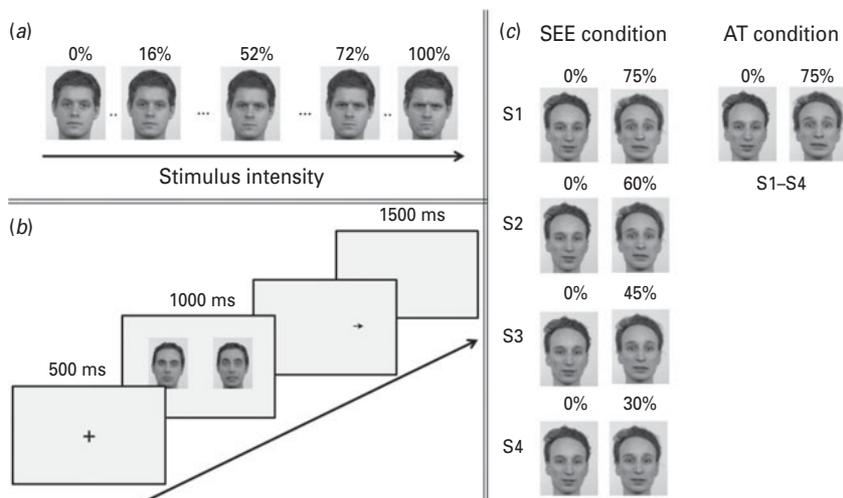


Fig. 1. (a) Examples of affective facial stimuli employed in the animated morph task. (b) Temporal trial structure for the training procedure. (c) Examples of stimuli employed in the training protocol. SEE, Sensitivity to emotional expressions; AT, attention training; S1–S4, sessions 1 to 4.

assessment in order to examine differences in facial emotion recognition between groups. To control for repetition effects in the animated morph task, half of the healthy CTLs were asked to additionally complete the post-training assessment 6 weeks following the baseline, paralleling the time span in the AVO group. Stimulus presentation and data collection were controlled by Presentation version 14.1 (Neurobehavioral Systems, USA) throughout all phases of the investigation. The experiment was run on a 15.4 inch (39 cm) WXGA wide TFT LCD notebook monitor and facial stimuli were presented at a viewing distance of about 50 cm at the center of the computer screen against a grey background.

Baseline

At baseline, participants were asked to complete a number of self-report measures including demographic information and questionnaire measures (BPAQ, PPI-R). They were then introduced to the animated morph task. Morphed images were presented for 500 ms, beginning with the neutral face that progressed into one of the six affective expressions. This procedure created the impression of an animated clip depicting the development of facial emotive expressions (see online Supplementary animated file). In order to increase task difficulty and to avoid a perfect correlation between time and emotional expression intensity, we occasionally repeated the same morph at random before presenting the next intensity morph level (Joormann & Gotlib, 2006). Therefore, a complete sequence consisted of 51 unique intensity levels but 71 image presentations. In our experiment, participants had to press a button as soon as they were able to

identify the emerging expression. The sequence was then immediately stopped, the face disappeared, and subjects were presented with a mask asking them to indicate the emotional expression they had just identified in a multiple-choice manner. Given a correct response, the intensity of emotional expression at the time of the button-press was averaged for every affective condition and included in statistical analysis. The experiment consisted of 72 trials with three model identities (different from those utilized in the training) each exhibiting six emotions, each with four repetitions in random order. All participants completed two practice trials before the experiment.

Training

AVOs were randomly assigned to either the attention (AT) or the sensitivity to emotional expressions (SEE) training condition. Participants as well as experimental investigators were blind to the experimental condition. The training sequence comprised four weekly sessions. During each trial of the training session, the participants were first presented with a fixation cross (500 ms), which indicated the beginning of a trial and was immediately followed by a bilateral presentation of a neutral and a fearful image of the same model identity (1 s). The fearful expression was always replaced by an arrow pointing to the left or the right which remained active until the participant indicated via button-press which direction the arrow was pointing to (Fig. 1b). The model identity, the position of the fearful cue and the arrow probe direction were pseudo-randomized across trials with no more than three identical sequential occurrences on each parameter. Each session consisted of 360 trials in total,

Table 1. Demographic and clinical sample characteristics

	AVOs (<i>n</i> =44)	CTLs (<i>n</i> =43)	Statistics
Age, years	35.32 (10.80)	29.34 (10.46)	$t_{85}=2.62, p=0.01$
BPAQ			
Physical aggression	25.25 (3.91)	19.20 (5.33)	$t_{85}=4.17, p<0.001$
Verbal aggression	15.00 (3.91)	14.30 (3.04)	$t_{85}=0.98, p>0.1$
Anger	16.23 (6.03)	13.09 (4.36)	$t_{85}=2.77, p<0.01$
Hostility	22.30 (7.36)	19.76 (5.60)	$t_{85}=1.79, p<0.1$
Total score	78.77 (20.26)	66.37 (14.72)	$t_{85}=3.25, p<0.01$
PPI-R			
Blame externalization	39.20 (6.61)	–	–
Rebellious non-conformity	61.18 (17.42)	–	–
Stress immunity	41.48 (6.47)	–	–
Social potency	46.05 (7.02)	–	–
Cold-heartedness	32.16 (7.97)	–	–
Machiavellian egocentricity	38.73 (6.84)	–	–
Carefree non-planfulness	30.02 (8.70)	–	–
Fearlessness	20.93 (5.06)	–	–
Insincere answering	42.77 (6.13)	–	–
Total score	309.75 (33.28)	–	–

AVOs, Antisocial violent offenders; CTLs, control participants; BPAQ, Buss–Perry Aggression Questionnaire; PPI-R, Psychopathic Personality Inventory–Revised.

Data are given as mean (standard deviation) of raw values.

with 120 distinct trial types (30 model identities \times 2 cue positions \times 2 arrow directions) and three repetitions. In the AT group, the training was performed with neutral and 75% fearful expressions throughout all four sessions. In the SEE training, the participants were trained with neutral and 75% fearful expressions only in the first session; the intensity of the fearful cue was successively decreased by 15% at every subsequent session, i.e. 60% intensity at the second, 45% at the third and 30% at the final session (Fig. 1c). In both training conditions the arrow always appeared in the eye region of the preceding emotional stimulus.

Post-training assessment

Two weeks following the last training session, the AVO group completed a second assessment of the morphing task conducted in the same manner as at baseline. In addition, 20 CTLs were reassessed on the animated morph task in order to control for repetition effects.

Results

Demographic and clinical data

A total of 44 AVOs and 43 CTLs were included in the final data analysis. All participants had completed secondary general school and thus did not differ with

regard to education, but AVO participants tended to be older and exhibited elevated BPAQ scores (Table 1).

For the interventional part of our study, separate analyses of variance (ANOVAs) were computed in order to investigate whether the subgroups (22 SEE training, 22 AT and 20 CTL participants) differed with regard to demographic or clinical measures. Results revealed that the groups did not differ with regard to age (SEE: mean=34.86 years, s.d.=11.09 years; AT: mean=35.77 years, s.d.=9.77 years; CTL: mean=34.00 years, s.d.=11.96 years; $F_{2,63}=0.13, p>0.1$). Significant differences between groups were evident only on the BPAQ total score (SEE: mean=78.23, s.d.=21.36; AT: mean=79.32, s.d.=19.59; CTL: mean=62.95, s.d.=14.80; $F_{2,63}=4.83, p<0.05$) and the subscales physical aggression (SEE: mean=25.04, s.d.=8.17; AT: mean=25.45, s.d.=8.17; CTL: mean=17.67, s.d.=5.36; $F_{2,63}=7.60, p<0.001$) and anger (SEE: mean=15.73, s.d.=6.45; AT: mean=16.72, s.d.=5.68; CTL: mean=12.30, s.d.=4.58; $F_{2,63}=3.49, p<0.05$), which were due to significant differences between the CTLs and both of the AVO subgroups as revealed by *post-hoc* analyses. The AVO subgroups did not differ on any of the BPAQ or PPI-R subscales (all p 's > 0.05).

Animated morph task at baseline

In order to investigate the emotion recognition deficits in AVOs, intensity levels at the time of the button-press

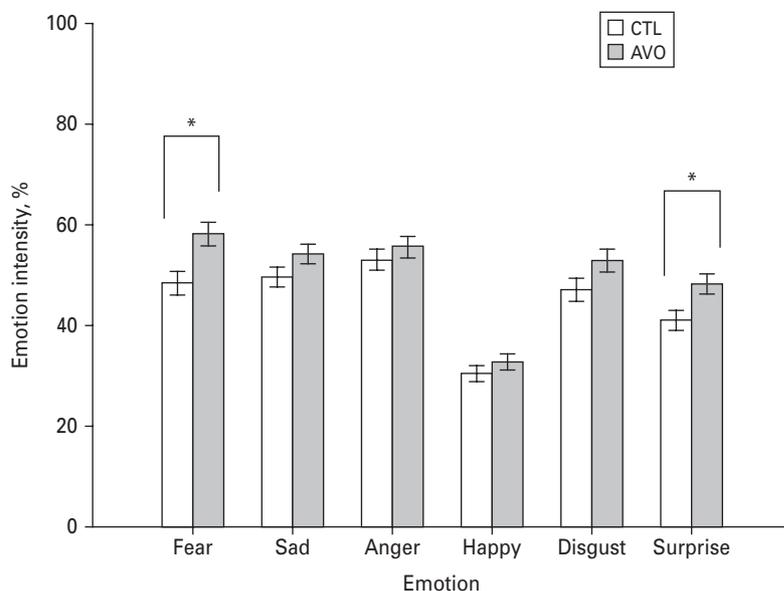


Fig. 2. Percentage of emotion intensity required to correctly detect the type of emotional expression by group. Values are means, with standard error of the mean represented by vertical bars. * Mean value was significantly different from that of the control group (CTL) ($p < 0.05$). AVO, Antisocial violent offender group.

for correct responses were analysed by calculating a 6 (emotion: happy, angry, fearful, sad, disgusted, surprised) \times 2 (group: AVO, CTL) repeated-measures ANOVA including age as a covariate (Fig. 2). A significant main effect of emotion ($F_{5,420} = 8.70$, $p < 0.001$) emerged and was further qualified by a significant emotion \times group interaction ($F_{5,420} = 2.29$, $p < 0.05$), whereas the main effect of group was significant only on a trend level ($F_{1,85} = 3.44$, $p < 0.1$). Separately computed t tests revealed that AVOs did not differ from CTLs in their identification of happy ($t_{85} = 1.08$, $p > 0.1$) or angry ($t_{85} = 0.88$, $p > 0.1$) faces. AVOs exhibited significantly impaired recognition of fearful ($t_{85} = 3.15$, $p < 0.01$) and surprised ($t_{85} = 2.64$, $p = 0.01$) expressions. The recognition deficits in AVOs regarding sad ($t_{85} = 1.71$, $p < 0.1$) and disgusted ($t_{85} = 1.73$, $p < 0.1$) expressions approached statistical significance on a trend level. These findings indicate that compared with CTLs, the AVOs exhibited an emotion recognition deficit that was most prominent for fearful and surprised expressions.

In order to rule out differential speed/accuracy trade-offs, we conducted an additional analysis of the error rates with a 6 (emotion: happy, angry, fearful, sad, disgusted, surprised) \times 2 (group: AVO, CTL) repeated-measures ANOVA. There was a significant main effect of emotion ($F_{5,420} = 32.09$, $p < 0.001$), indicating that both groups made more errors for fearful, disgusted and surprised than for angry, sad or happy faces. There was no significant main effect of group ($F_{1,85} = 0.56$, $p > 0.1$) and no significant interaction ($F_{5,420} = 1.08$, $p > 0.1$). These findings indicate that

emotion recognition deficits in the AVO group cannot be explained by a speed/accuracy trade-off. As evident in the confusion matrices provided in online Supplementary Table S1, disgust was frequently confused with anger and fear with surprise, a characteristic pattern that was similar between groups and is in accordance with previous literature (Fairchild *et al.* 2009).

Post-training assessment

In order to investigate the training and repetition effects, we computed difference scores between the correct recognition performance for each emotional expression on the animated morph task at baseline and at the second assessment ($\text{mean}_{\Delta} = \text{mean}_{\text{baseline}} - \text{mean}_{\text{post}}$), with positive scores indicating an improvement. The difference scores were then analysed with a 6 (emotion: happy, angry, fearful, sad, disgusted, surprised) \times 3 (group: SEE training, AT, CTL) repeated-measures ANOVA (Fig. 3). The results indicated a significant main effect of emotion ($F_{5,305} = 2.56$, $p < 0.05$) and group ($F_{2,61} = 4.18$, $p < 0.05$), whereas the emotion \times group interaction was non-significant ($F_{10,305} = 1.09$, $p > 0.1$). *Post-hoc* analyses revealed that the group effect was due to significantly better performance in the SEE training compared with the AT group ($p < 0.05$) and a trend toward significantly better performance than the CTL group ($p < 0.1$), whereas the CTL and AT groups did not differ in their performance. These results indicate that only the SEE group exhibited an improvement in emotion recognition.

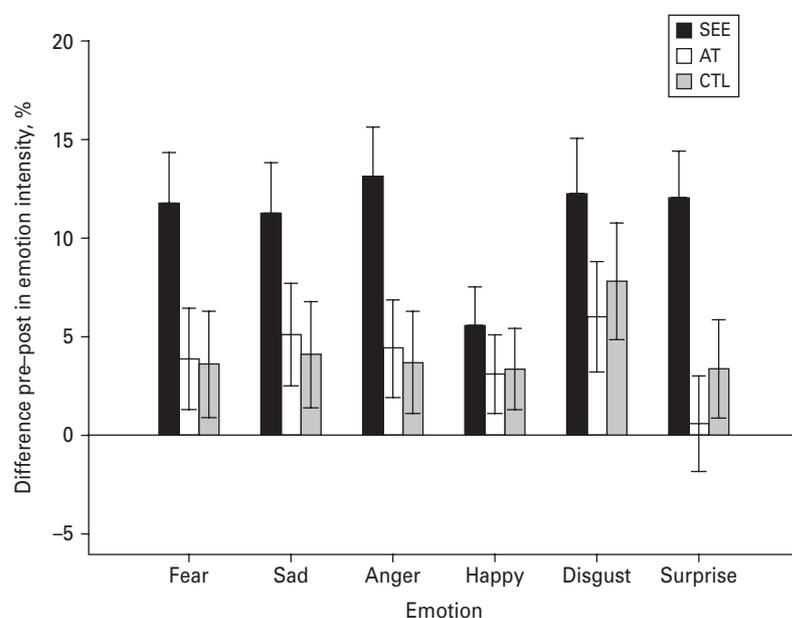


Fig. 3. Change in the percentage of emotion intensity required to correctly detect the type of emotional expression by group. The change was calculated by difference scores between the correct recognition performance for each emotional expression on the morphing task at baseline and at the second assessment ($\text{mean}_{\Delta} = \text{mean}_{\text{baseline}} - \text{mean}_{\text{post}}$), with positive scores indicating an improvement. Values are means, with standard error of the mean represented by vertical bars. SEE, Sensitivity to emotional expressions; AT, attention training; CTL, control group.

An additional analysis of the pre-post error rates with a 6 (emotion: happy, angry, fearful, sad, disgusted, surprised) \times 2 (group: AVO, CTL) repeated-measures ANOVA revealed no significant main effect of emotion ($F_{5,305} = 0.89$, $p > 0.1$) or group ($F_{1,61} = 1.78$, $p > 0.1$) and no significant interaction ($F_{5,305} = 0.23$, $p > 0.1$).

Discussion

The present study investigated facial affect recognition deficits in incarcerated violent offenders with psychopathic traits relative to control participants. Consistent with previous studies, our data revealed that aggressive individuals exhibited the most pronounced impairment in the recognition of a developing fearful emotion (Marsh & Blair, 2008). In addition, antisocial individuals also exhibited a decreased sensitivity to faces displaying surprised expressions, a finding that most probably reflects the difficulty to discriminate fear and surprise at an early perceptual level (Young *et al.* 1997). Neither the recognition performance of the remaining emotions nor the error patterns resulted in significant differences between groups. However, inmates displayed a delayed detection of sad and disgusted expressions on a statistical trend level, which may suggest that deficits in affect recognition are not restricted to specific emotions. This finding is in accordance with the results of a more recently

published meta-analysis indicating that emotion recognition impairments in psychopathy are pervasive across emotions (Dawel *et al.* 2012), and also consistent with recent theorizing about the broader role of the amygdala in emotion processing (Adolphs, 2010).

With regard to the second aim of the present study, we were able to demonstrate that a brief computerized training protocol was sufficient to significantly improve recognition of facial affect. More specifically, we demonstrated that a modified implicit ABM training with step-wise intensity reductions of the fearful facial cues not only led to a lower threshold in the recognition of fear, but also improved the sensitivity to all affective expressions. This unexpected finding suggests that individuals successively learned to detect and interpret even subtle changes in fearful faces and that the acquisition of these 'emotion reading' skills generalizes to other affective expressions. Furthermore, our data revealed that the training effect appeared only in the SEE training but not in the AT condition, indicating that directing the attentional focus to salient parts of a facial expression alone did not alleviate the perceptual deficits evident in aggressive individuals.

It could be argued that the latter finding contradicts previous evidence that demonstrated a transient alteration of the fear recognition impairment in child psychopathy after instructing the participants to focus on the eye region of prototypical fearful faces (Dadds *et al.* 2006, 2008). The eye region is crucial for the

portrayal of emotions and a lack of attention to this relevant face region will definitely result in poor affect recognition. However, training to focus on the eye region of prototypical full-blown face portraits may not translate to improvements in everyday social interactions that require deciphering more subtle signs of a developing affect in other people's eyes. To improve perceptual sensitivity to less intense affective cues in facial expressions, the individual has to learn how salient features of a neutral face convert into an emotional expression. Thus, the present study extends the intriguing results of the previous work by establishing a training tool that uses the implicit direction of attention to salient face areas and manipulates the affective intensity of the stimulus material in order to sensitize for early aspects of a developing emotion.

It must be noted that attempts to target emotion recognition deficits in other clinical groups have been documented in several recent studies. Social-cognitive remediation or empathy trainings have recently been investigated in severe mental disorders, including autistic and schizophrenic patients (Russell *et al.* 2006; Marsh *et al.* 2010, 2012), typically under the employment of the Ekman micro-expression training tool (METT, www.paulekman.com), which uses a series of training videos to draw attention to important distinguishing features (eyes, nose, mouth) of affective facial expressions. In schizophrenic patients, METT has been shown to improve emotion recognition accuracy and to have an impact on visual scanning for novel facial stimuli (Marsh *et al.* 2010, 2012). Notably, Dadds *et al.* (2012) recently investigated the efficacy of an empathic-emotion recognition training (ERT) in a population of children with complex conduct problems and reported that particularly individuals scoring high on callous-unemotional traits exhibited improvements in problem behaviors. However, the authors failed to confirm that ERT was the effective component of the training, as they found no evidence for an improvement in facial emotion recognition. Only recently, a study by Penton-Voak *et al.* (2013) demonstrated that shifting the categorical boundary between angry and happy facial expressions in a morphed continuum of images through biased feedback was sufficient to encourage the perception of happiness over anger in ambiguous expressions. Most interestingly, the authors could show that their interpretation bias modification training resulted in a decrease in self-reported as well as staff-rated aggressive behavior in high-risk youth (Penton-Voak *et al.* 2013).

Our study, however, is not only the first to date to investigate the utility of an emotion recognition training in aggressive adults, but also extends previous findings with regard to several aspects. First, this investigation is unique in its methodological approach

in that it is strictly implicit, unlike the training methods employed in previous studies that typically involve explicit instructions to attend to the central features of a face. Furthermore, we extend the above-mentioned findings in showing that attention is not the only component that can be utilized to enhance emotion recognition; the manipulation of stimulus intensity appears to be a crucial element contributing to the pronounced improvement of emotion recognition. Third, the assessment of the dynamic onset of an emotional expression instead of the recognition of static full-blown emotions as an outcome variable may represent a more valid tool to capture changes in the ability to detect subtle emotional cues (Hastings *et al.* 2008; Pham & Philippot, 2010). Finally, we demonstrated the training effect to be detectable in a different task modality, to hold for at least 2 weeks and not to be driven by mere task repetition effects.

Several limitations of our study should be considered. First, the training was designed to implicitly direct participants' attention to the eye region of the presented faces. However, in the present study we did not use eye-tracking equipment to measure actual gaze behavior during the training as well as in the morphing task pre-/post-intervention. Thus, future studies should investigate whether this type of training also has an impact on visual scanning of emotional faces. In this regard, there is also a need to clarify whether a free gaze condition in the SEE training (i.e. attention is not shifted to the eye region) would also be appropriate to improve emotion recognition. Further, it is unclear whether our SEE training produces general face awareness and face recognition benefits that are not restricted to an improvement in recognition of emotional expressions. This could be implemented, for instance, by including a training condition in which faces are morphed along the male-female gender boundaries and the attention of the subject is systematically directed toward an increasingly androgynous face. If generalized face awareness underlies the training benefits, the gender control condition should likewise be expected to result in enhanced expression recognition in the morphing task. A second limitation of the present work is that we included only incarcerated male offenders with psychopathic tendencies and it remains an open question whether the present findings can be generalized to female psychopaths or non-psychopathic antisocial populations. In this context, psychopathy also needs to be assessed more cautiously by administering diagnostic interviews (e.g. the Psychopathy Checklist; Hare, 2003) as well as instruments allowing us to elucidate what kind of psychopathic traits may be particularly associated with the recognition deficits (e.g. primary or secondary psychopathy, high

callous-unemotional traits). Follow-up studies should address this issue by considering groups high and low in this trait, thus allowing us to evaluate the efficacy of the training in various subsamples. In more general terms, it must be mentioned that the present findings are preliminary and based on a small sample of violent inmates. Future replication of these findings in larger cohorts is important for increasing confidence in the observed pattern of findings in order to draw valid implications for treatment. Issues concerning the durability of the intervention effects as well as the neural mediators of these alterations should also be subject to future studies. However, the main question that needs to be addressed by future studies is whether an emotion sensitivity training is associated with an improvement in relevant behavioral outcomes. Employment of frustration tasks, the assessment of physiological activation patterns in response to challenging social situations, or blinded peer ratings in non-incarcerated samples could be utilized to operationalize aggressive behavioral tendencies as relevant outcome measures.

In summary, this study adds to the body of evidence that antisocial behavior is associated with impaired recognition of facial affect. We demonstrated that violent offenders with psychopathic traits require significantly higher intensities of a developing facial emotion, especially of fear, to accurately categorize the expression in animated film clips. Further, the present study shows for the first time that this deficit can be addressed by an implicit training that directs the attentional focus to salient regions of a face and gradually decreases the intensity of the emotional expression. Future studies should focus on the potential of this intervention to effectively increase empathy skills and inhibit violent behavior in antisocial individuals.

Supplementary material

For supplementary material accompanying this paper visit <http://dx.doi.org/10.1017/S0033291713001517>.

Declaration of Interest

None.

References

- Adolphs R** (2010). What does the amygdala contribute to social cognition? *Annals of the New York Academy of Sciences* **1191**, 42–61.
- Adolphs R, Gosselin F, Buchanan TW, Tranel D, Schyns P, Damasio AR** (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature* **433**, 68–72.
- Bar-Haim Y** (2010). Research review: attention bias modification (ABM): a novel treatment for anxiety disorders. *Journal of Child Psychology and Psychiatry* **51**, 859–870.
- Blair R, Mitchell D, Peschardt K, Colledge E, Leonard R, Shine J, Murray L, Perrett D** (2004). Reduced sensitivity to others' fearful expressions in psychopathic individuals. *Personality and Individual Differences* **37**, 1111–1122.
- Blair RJR** (2001). Neurocognitive models of aggression, the antisocial personality disorders, and psychopathy. *Journal of Neurology, Neurosurgery and Psychiatry* **71**, 727–731.
- Buss AH, Perry M** (1992). The Aggression Questionnaire. *Journal of Personality and Social Psychology* **63**, 452–459.
- Dadds MR, Cauchi AJ, Wimalaweera S, Hawes DJ, Brennan J** (2012). Outcomes, moderators, and mediators of empathic-emotion recognition training for complex conduct problems in childhood. *Psychiatry Research* **199**, 201–207.
- Dadds M, El Masry Y, Wimalaweera S, Guastella A** (2008). Reduced eye gaze explains 'fear blindness' in childhood psychopathic traits. *Journal of the American Academy of Child and Adolescent Psychiatry* **47**, 455–463.
- Dadds M, Perry Y, Hawes D, Merz S, Riddell A, Haines D, Solak E, Abeygunawardane A** (2006). Look at the eyes: fear recognition in child psychopathy. *British Journal of Psychiatry* **189**, 180–181.
- Dawel A, O'Kearney R, McKone E, Palermo R** (2012). Not just fear and sadness: meta-analytic evidence of pervasive emotion recognition deficits for facial and vocal expressions in psychopathy. *Neuroscience and Biobehavioral Reviews* **36**, 2288–2304.
- Dolan M, Fullam R** (2006). Face affect recognition deficits in personality-disordered offenders: association with psychopathy. *Psychological Medicine* **36**, 1563–1569.
- Eisenbarth H, Alpers GW** (2008). *Psychopathic Personality Inventory-Revised, German Version (PPI-R)*. Hogrefe: Göttingen.
- Eisenbarth H, Alpers GW, Segrè D, Calogero A, Angrilli A** (2008). Categorization and evaluation of emotional faces in psychopathic women. *Psychiatry Research* **159**, 189–195.
- Fairchild G, Stobbe Y, Van Goozen SHM, Calder AJ, Goodyer IM** (2010). Facial expression recognition, fear conditioning, and startle modulation in female subjects with conduct disorder. *Biological Psychiatry* **68**, 272–279.
- Fairchild G, Van Goozen SHM, Calder AJ, Stollery SJ, Goodyer IM** (2009). Deficits in facial expression recognition in male adolescents with early-onset or adolescence-onset conduct disorder. *Journal of Child Psychology and Psychiatry* **50**, 627–636.
- Hare RD** (2003). *The Psychopathy Checklist – Revised*, 2nd edn. Multi-Health Systems: Toronto.
- Harms MB, Martin A, Wallace GL** (2010). Facial emotion recognition in autism spectrum disorders: a review of behavioral and neuroimaging studies. *Neuropsychology Review* **20**, 290–322.

- Hastings ME, Tangney JP, Stuewig J** (2008). Psychopathy and identification of facial expressions of emotion. *Personality and Individual Differences* **44**, 1474–1483.
- Herzberg PY** (2003). Faktorstruktur Gütekriterien und Konstruktvalidität der deutschen Übersetzung des Aggressionsfragebogens von Buss und Perry [Psychometric evaluation and validity of the German translation of the Aggression Questionnaire by Buss and Perry]. *Zeitschrift für Differentielle und Diagnostische Psychologie* **24**, 311–323.
- Jones A, Laurens K, Herba C, Barker G, Viding E** (2009). Amygdala hypoactivity to fearful faces in boys with conduct problems and callous-unemotional traits. *American Journal of Psychiatry* **166**, 95–102.
- Joormann J, Gotlib IH** (2006). Is this happiness I see? Biases in the identification of emotional facial expressions in depression and social phobia. *Journal of Abnormal Psychology* **115**, 705–713.
- Kohler CG, Walker JB, Martin EA, Healey KM, Moberg PJ** (2010). Facial emotion perception in schizophrenia: a meta-analytic review. *Schizophrenia Bulletin* **36**, 1009–1019.
- Kosson DS, Suchy Y, Mayer AR, Libby J** (2002). Facial affect recognition in criminal psychopaths. *Emotion* **2**, 398–411.
- Langner O, Dotsch R, Bijlstra G, Wigboldus DHJ, Hawk ST, Van Knippenberg A** (2010). Presentation and validation of the Radboud Faces Database. *Cognition and Emotion* **24**, 1377–1388.
- Lilienfeld SO, Widows MR** (2005). *Psychopathic Personality Inventory-Revised (PPI-R): Professional Manual*. Psychological Assessment Resources: Lutz, FL.
- MacLeod C, Mathews A, Tata P** (1986). Attentional bias in emotional disorders. *Journal of Abnormal Psychology* **95**, 15–20.
- Marsh AA, Blair R** (2008). Deficits in facial affect recognition among antisocial populations: a meta-analysis. *Neuroscience and Biobehavioral Reviews* **32**, 454–465.
- Marsh PJ, Green MJ, Russell TA, McGuire J, Harris A, Coltheart M** (2010). Remediation of facial emotion recognition in schizophrenia: functional predictors, generalizability, and durability. *American Journal of Psychiatric Rehabilitation* **13**, 143–170.
- Marsh PJ, Luckett G, Russell T, Coltheart M, Green MJ** (2012). Effects of facial emotion recognition remediation on visual scanning of novel face stimuli. *Schizophrenia Research* **141**, 234–240.
- Montagne B, van Honk J, Kessels RP, Frigerio E, Burt M, van Zandvoort MJ, Perrett DI, de Haan EH** (2005). Reduced efficiency in recognising fear in subjects scoring high on psychopathic personality characteristics. *Personality and Individual Differences* **38**, 5–11.
- Penton-Voak IS, Thomas J, Gagel SH, McMurran M, McDonald S, Munafò MR** (2013). Increasing recognition of happiness in ambiguous facial expressions reduces anger and aggressive behavior. *Psychological Science* **24**, 688–697.
- Pham TH, Philippot P** (2010). Decoding of facial expression of emotion in criminal psychopaths. *Journal of Personality Disorders* **24**, 445–459.
- Raine A, Lee L, Yang Y, Colletti P** (2010). Neurodevelopmental marker for limbic maldevelopment in antisocial personality disorder and psychopathy. *British Journal of Psychiatry* **197**, 186–192.
- Russell TA, Chu E, Phillips ML** (2006). A pilot study to investigate the effectiveness of emotion recognition remediation in schizophrenia using the micro-expression training tool. *British Journal of Clinical Psychology* **45**, 579–583.
- Sato W, Uono S, Matsuura N, Toichi M** (2009). Misrecognition of facial expressions in delinquents. *Child and Adolescent Psychiatry and Mental Health* **3**, 27–34.
- Schönenberg M, Louis K, Mayer S, Jusyte A** (2013). Impaired identification of threat-related social information in male delinquents with antisocial personality disorder. *Journal of Personality Disorders*. Published online 15 April 2013. doi:10.1521/pedi_2013_27_100.
- Sebastian CL, McCrory EJ, Cecil CA, Lockwood PL, De Brito SA, Fontaine NM, Viding E** (2012). Neural responses to affective and cognitive theory of mind in children with conduct problems and varying levels of callous-unemotional traits. *Archives of General Psychiatry* **69**, 814–822.
- Staugaard SR** (2010). Threatening faces and social anxiety: a literature review. *Clinical Psychology Review* **30**, 669–690.
- Young AW, Rowland D, Calder AJ, Etcoff NL** (1997). Facial expression megamix: tests of dimensional and category accounts of emotion recognition. *Cognition* **63**, 271–313.