AGGRESSIVE BEHAVIOR
Volume 9999, pages 1–12 (2017)

Generalized Hostile Interpretation Bias Regarding Facial Expressions: Characteristic of Pathological Aggressive Behavior

Danique Smeijers[1,4]*, Mike Rinck[2], Erik Bulten[3], Thom van den Heuvel[1], and Robbert-Jan Verkes[1,3,4]

[1]Department of Psychiatry, Radboud University Medical Center, Nijmegen, The Netherlands
[2]Behavioural Science Institute, Radboud University Nijmegen, Nijmegen, The Netherlands
[3]Pompestichting, Nijmegen, The Netherlands
[4]Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands

Individuals with aggression regulation disorders tend to attribute hostility to others in socially ambiguous situations. Previous research suggests that this "hostile attribution bias" is a powerful cause of aggression. Facial expressions form important cues in the appreciation of others' intentions. Furthermore, accurate processing of facial expressions is fundamental to normal socialization. However, research on interpretation biases in facial affect is limited. It is asserted that a hostile interpretation bias (HIB) is likely to be displayed by individuals with an antisocial (ASPD) and borderline personality disorder (BPD) and probably also with an intermittent explosive disorder (IED). However, there is little knowledge to what extent this bias is displayed by each of these patient groups. The present study investigated whether a HIB regarding emotional facial expressions was displayed by forensic psychiatric outpatients (FPOs) and whether it was associated with ASPD and BPD in general or, more specifically, with a disposition to react with pathological aggression. Participants of five different groups were recruited: FPOs with ASPD, BPD, or IED, non-forensic patients with BPD (nFPOs-BPD), and healthy, non-aggressive controls (HCs). Results suggest that solely FPOs with ASPD, BPD, or IED exhibit a HIB regarding emotional facial expressions. Moreover, this bias was associated with type and severity of aggression, trait aggression, and cognitive distortions. The results suggest that a HIB regarding facial expressions is an important characteristic of pathological aggressive behavior. Interventions that modify the HIB might help to reduce the recurrence of aggression. Aggr. Behav. 9999:1–12, 2017. **© 2017 Wiley Periodicals, Inc.**

Keywords: aggression; bias; facial-expressions; hostility

INTRODUCTION

Facial Expressions

An important aspect of social interaction is facial expression. Facial expressions carry a broad range of socially relevant information: They give information regarding the internal states of the sender and they provide important cues for our impressions of others (Moriya, Tanno, & Sugiura, 2012; Schonenberg & Jusyte, 2014). Accurate processing of facial expressions is critical to normal socialization and social interaction (Moriya et al., 2012). However, in everyday life, situations are not always straightforward and often consist of ambiguous information. Likewise, people do not necessarily express prototypical expressions; subtle facial expressions are often displayed instead (Moriya et al., 2012). This renders it even more difficult to process facial expressions correctly. Furthermore, associated with psychopathology, some individuals lack the ability to process facial affect accurately. For instance, antisocial personality and psychopathic traits have been associated with deficits in the recognition of facial affect (Dawel, O'Kearney, McKone, & Palermo, 2012; Dolan & Fullam, 2006; Marsh & Blair, 2008). However, incorrect recognition may not solely be based on impairments in emotion recognition; it might also

Contract grant sponsor: NWO/National Initiative Brain and Cognition NIHC; contract grant number: NWO-056-24-011.
Conflicts of interest: On behalf of all authors, the corresponding author states that there is no conflict of interest.
*Correspondence to: Danique Smeijers, Department of Psychiatry (966), Radboud University Medical Centre, PO Box 9101, 6500 HB, Nijmegen, The Netherlands.
E-mail: daniquesmeijers@gmail.com

Received 7 June 2016; Revised 22 November 2016; Accepted 1 December 2016

DOI: 10.1002/ab.21697
Published online XX Month Year in Wiley Online Library (wileyonlinelibrary.com).

arise from an a-priori tendency to interpret facial expression in a certain way, regardless of the ability to recognize the expression correctly. In this case, we speak of an interpretation bias.

Hostility Biases

Apart from the processing of facial expression of others, aggressive behavior has been suggested to occur after making a hostile attribution that "the self" has been threatened (Dodge, 2006). This tendency of aggressive individuals to perceive or attribute hostile intent to others is often referred to as "hostile attribution bias" (Nasby, Hayden, & DePaulo, 1980). Reactive aggressive behavior is thought to be associated with this hostile attribution bias, in children as well as in adults; higher levels of this bias were associated with higher levels of aggressive behavior (Bailey & Ostrov, 2007; Chen, Coccaro, & Jacobson, 2012; Crick, 1995; Crick & Dodge, 1996; Crick, Grotpeter, & Bigbee, 2002; De Castro, Veerman, Koops, Bosch, & Monshouwer, 2002; Dodge, 2006). Reactive aggression is defined as an impulsive, angry, or defensive response to threat, provocation, or frustration (Crick & Dodge, 1996), and occurs in response to particular social situations. Furthermore, these studies suggest that aggressive individuals tend to attribute hostility to others in socially ambiguous situations. This attribution of hostility can have detrimental effects, as the perception of aggressive intent in others is a powerful cause of anger and aggressive behavior (Epstein & Taylor, 1967). Moreover, hostile attribution biases not only cause and predict aggressive acts (Dodge, Bates, & Pettit, 1990), they also contribute to the maintenance of aggression. When attributing hostile intents to others, one is more likely to act aggressively, which in turn causes others to respond more aggressively, which will further strengthen the person's hostile view of others (Crick & Dodge, 1996; Helfritz-Sinville & Stanford, 2014).

Previous studies, which revealed the existence of hostile attribution biases, solely made use of vignettes, using videos or written stories. The vignettes describe hypothetical situations in which someone is provoked by a peer who is acting ambiguously. Participants are asked to indicate the intention of the peer. Research showed that a hostile attribution bias occurs in these ambiguous situations. With regard to this aspect of hostile biases, vignettes are suitable materials. However, as stated earlier, facial expressions play an important role in social situations and may also induce aggressive intentions and behavior when not processed or interpreted accurately. It is possible that this hostile bias not only occurs in the attribution of others' intentions in ambiguous situations, but also with respect to the interpretation of ambiguous facial expressions. However, the method of vignettes omits the possible role of facial expression in a hostile interpretation bias (HIB).

Until now, only a few studies have investigated interpretation biases regarding facial expressions. Hoaken, Allaby, and Earle (2007) have observed that violent offenders, compared to non-violent offenders and healthy controls, were likely to interpret a neutral face less often as sad and more frequently as disgusted. Another study revealed that an attributional bias towards "negative facial affect" was associated with aggressiveness in institutionalized boys (Nasby et al., 1980). Burt, Mikolajewski, and Larson (2009) showed that trait aggression was moderately associated with hostile perceptions of others' neutral facial expressions. Schonenberg and Jusyte (2014) recently investigated the HIB toward ambiguous facial cues in antisocial violent offenders. They used pictures of emotional faces (angry, happy, and fearful) and morphed them with each other, creating three dimensions: happy-fearful, happy-angry, and fearful-angry. They created five distinct intensity levels. Their results suggest that antisocial violent offenders infer hostile intent from the angry-happy and -fearful morphs. The authors conclude that aggressive individuals interpreted ambiguous facial cues as hostile and showed a tendency to overrate the perceived intensity of anger. However, the morphed pictures all consisted of two emotions shown simultaneously. Even though these stimuli are ambivalent, in everyday life, these emotions are rarely displayed together.

Previous studies of biases assume that rating a picture as angry suggests that the person also interprets it as hostile. This notion suggests that anger and hostility are interchangeable. However, "angry" and "hostile" differ in meaning (Eckhardt, Norlander, & Deffenbacher, 2004). Anger can be referred to as an emotional state and constellation of specific uncomfortable subjective experiences and associated cognitions, whereas hostility can be described as an attitude that involves dislike and negative evaluation of others (Eckhardt et al., 2004). To our knowledge, it is unknown whether individuals interpret only an angry face as hostile. It could be that other emotions such as disgust are also experienced as hostile. In addition, none of the previous studies asked participants to indicate whether they experienced a certain picture as hostile.

Hostility Biases and Psychopathology

In addition to the general population samples tested in previous research, a hostile attribution/interpretation bias was also found in a few studies of populations displaying pathological forms of aggressive behavior (Lobbestael, Cima, & Arntz, 2013; Schonenberg & Jusyte, 2014). Pathological aggression, here, is defined

as aggressive behavior which is disproportionate to the provocation (Siever, 2008). Pathological aggression is a characteristic feature of antisocial personality disorder (ASPD), borderline personality disorder (BPD) and intermittent explosive disorder (IED). These two latter disorders were less intensively studied regarding hostile attribution/interpretation biases, although they seem to be as relevant as ASPD.

Borderline personality disorder (BPD) is characterized by inappropriate and intense anger or difficulty controlling anger (APA, 2000). Patients diagnosed with BPD are highly likely to display maladaptive cognitive processes (Baer, Peters, Eisenlohr-Moul, Geiger, & Sauer, 2012). For instance, borderline patients tend to believe that other individuals are more hostile, aggressive, untrustworthy, and dangerous (Arntz, Weertman, & Salet, 2011; Baer et al., 2012; Fertuck, Grinband, & Stanley, 2013). Furthermore, associations have been found with attentional biases towards negative emotional stimuli and with rejection- and anger-related interpretation of ambiguous social situations (Baer et al., 2012; Lobbestael & McNally, 2015). With respect to facial expressions, borderline patients are thought to interpret neutral or ambiguous facial expressions more negatively and to exhibit deficits in the recognition of emotional facial expressions, in particular regarding faces displaying anger and disgust (Daros, Zakzanis, & Ruocco, 2013). Previous literature seems to suggest that BPD is as much associated with aggression and hostile interpretation biases as ASPD. This raises the question whether a HIB is a characteristic of ASPD and BPD in general.

When recurrent episodes of failure to resist aggressive impulses are the main characteristic of a patient, without clear evidence for a personality disorder, IED may be the most appropriate diagnosis (APA, 2000). Even though this disorder is not as persistent as personality disorders, its behavioral patterns may have as many detrimental consequences. IED has not been related to cognitive and hostility biases as frequently as ASPD and BPD. Nevertheless, one of the few studies regarding this topic showed that individuals with IED exhibit deficits in emotional intelligence (Coccaro, Solis, Fanning, & Lee, 2015). It is assumed that emotional intelligence is related to the ability to understand and recognize emotional information of oneself and others, and to use this information to guide thoughts, actions and coping mechanisms (Coccaro et al., 2015; Mayer & Salovey, 1993). Furthermore, this deficit was found to be associated with hostile cognitions as well as hostile attributions to others' intentions in socially ambiguous situations. Coccaro et al. (2015) suggest that a deficit in emotional intelligence facilitates the tendency to make hostile attributions about others' intentions. These results suggest that IED is as likely to be associated with a HIB as ASPD and BPD are.

According to previous literature a HIB is associated with reactive aggression and is likely to be displayed by individuals diagnosed with ASPD as well as BPD, and probably also those with IED. However, to our knowledge, there are no studies regarding hostility biases in which these psychiatric disorders were investigated simultaneously. Unraveling whether hostility biases are associated with general or pathological aggression or specific psychiatric disorders might be of particular importance for clinical practice. Treatment for aggression is based on whether or not there is a personality disorder. One variant generally consists of developing skills to control anger, often used in less profound disorders. Cognitive therapeutic interventions, on top of skill training, may be used in treatment variants often used among patients with personality disorders, for instance to alter cognitive distortions. As a HIB may be a powerful cause and maintaining factor of aggressive behavior, it may have to be reduced/changed in interventions for even less persistent and profound disorders, such as IED, as well. Until now, there is little knowledge about the extent to which this bias is specifically displayed by patients with ASPD, BPD, or IED.

Current Study

The aims of the present study, therefore, were twofold. First, we aimed to validate a newly developed test to investigate the presence of a HIB with respect to facial expressions. This HIB test consisted of a computer task in which facial pictures displaying emotions of various intensities were presented. The pictures displayed the following emotions: anger, happiness, fear, and disgust. To create different intensities, pictures were morphed with neutral pictures in order to avoid images showing multiple emotions simultaneously and to let the expressions be more similar to expressions displayed in daily life. Although the morphed pictures ranged from neutral to full emotion, they exhibited a certain degree of ambiguity; the neutral expression was in all models displayed with mouth closed whereas the emotional pictures where displayed with mouth open. This difference in mouth position resulted in pictures showing ambiguous expressions. Moreover, instead of asking participants to indicate which emotion was shown or to indicate the intensity of the emotion, participants were asked to indicate whether the picture looked hostile or not. Furthermore, previous research mainly focused on the link between hostile attribution/interpretation bias and reactive versus proactive aggression. To be able to discover a possible association between HIB and type and severity of aggression and

conscious cognitive distortions, a number of questionnaires were administered.

The second aim was to investigate whether a HIB regarding emotional facial expressions is specifically associated with ASPD, or with both ASPD and BPD or, more specifically, with pathological aggressive behavior. Participants from five different groups were recruited: forensic psychiatric outpatients (FPOs) diagnosed with ASPD, BPD, or IED, non-forensic borderline patients with BPD (nFPOs-BPD), and healthy, non-aggressive controls (HCs). It was predicted that all groups of FPOs would show a HIB. Specifically, in line with the work of Schonenberg and Jusyte (2014), we predicted that FPOs would interpret low intensity, ambiguous facial expressions more often as hostile than HCs would. In addition, it was predicted that nFPOs-BPD would also exhibit a greater HIB towards facial expressions than HCs.

METHODS

Participants

One-hundred-forty-two FPOs, 23 nFPOs-BPD of a general mental health institute and 47 HCs were recruited for the study. The FPOs were recruited from a series of patients admitted to "Kairos, Pompesticht-ing", an outpatient clinic for forensic psychiatry at Nijmegen, The Netherlands. The nFPOs- BPD were recruited at the department of psychiatry of Radboud University Medical Center, Nijmegen, The Netherlands. They were referred to a dialectical behavior therapy, which is a comprehensive and evidence-based treatment for BPD (Linehan and Read, 2013). They were diagnosed with BPD by trained clinicians at the department of psychiatry. The HCs were recruited via online postings on the clinic's website.

To exclude participants with a diagnosis of a lifetime bipolar disorder, psychosis, current major depression, or current severe addiction, all participants were screened with the MINI International Neuropsychiatric Interview (M.I.N.I; Sheehan et al., 1998; Van Vliet & De Beurs, 2007). None of the nFPOs-BPD and HCs, and only two of the FPOs had to be excluded due to a current major depression.

The FPOs were screened by trained clinicians with the Structured Clinical Interview for DSM-IV axis II personality disorders (SCID-II; Weertman, Arntz, & Kerkhofs, 2000) and the Research Criteria set for Intermittent Explosive Disorder (IED-R; Coccaro, Kavoussi, Berman, & Lish, 1998). For the purpose of the current study, solely FPOs diagnosed with either ASPD, BPD, or IED were included. 33 FPOs had to be excluded due to co-occurrence of these disorders. Therefore, the final current sample of FPOs consisted of 40 patients diagnosed with ASPD, 30 with BPD, and 37 with IED without ASPD or BPD. All FPOs were referred to the outpatient clinic because of aggression regulation problems. In the present study, 82 patients were referred voluntarily and 25 obligatory.

Both the FPOs and the nFPOs- BPD participated before the start of their treatment. After receiving information about the nature of the study, participants signed a consent form. Demographic information is provided in Table I. In total, it took participants approximately 50 min to complete the assessments. All participants were compensated for their time with an appropriate monetary reward. Moreover, the current study was approved by the Research Ethics Committee, CMO region Arnhem-Nijmegen, The Netherlands.

MATERIALS

Questionnaires

The social dysfunction and aggression scale (SDAS; Wistedt et al., 1990) is an observer-scale that measures the severity of actual aggressive behavior. It consists of nine items measuring outward aggression and two items measuring inward aggression. Items have to be scored on a 4-point scale with 0 = not present and 4 = severely to extremely present as extremes. The SDAS has adequate observer reliability (Cronbach's Alpha = .79; Wistedt et al., 1990). In the current study, due to lack of observers, the SDAS was used as self-report. Participants had to rate their aggressive behavior over a period of three months. Moreover, only FPOs with a total SDAS score of five points or higher were included in the study.

The reactive proactive questionnaire (RPQ; Cima, Raine, Meesters, & Popma, 2013; Raine et al., 2006) is a 23-item self-report questionnaire to measure reactive and proactive aggression. The reactive subscale consists of 11 items whereas the proactive subscale consists of 12 items. The items are rated 0 (never), 1 (sometimes), or 2 (often). The Dutch translation has good internal consistency (Cronbach's Alpha = .91) and adequate convergent (all $r < .16$), criterion (delinquents from prison and forensic mental health scored higher than non-offenders) and construct validity (violent offenders show more proactive aggression than non-offenders, $P < .001$; Cima et al., 2013). In the current study the internal consistency has also proven to be good (Cronbach's Alpha = .92).

The aggression questionnaire (AQ; Buss & Perry, 1992) is a self-report questionnaire to assess an overall trait of aggression. It consists of 29 items which are divided into four subscales: physical aggression, verbal aggression, anger, and hostility. The items are scored on a 5-point Likert scale (1 = extremely unlike me to

TABLE I. Demographic Information and Mean Scores and Standard Deviations on Questionnaires

Questionnaire	FPOs ASPD	FPOs IED	FPOs BPD	nFPOs BPD	HCs
Age	$M = 34.27$	$M = 40.08$	$M = 36.30$	$M = 34.83$	$M = 38.06$
	(SD = 9.31)	*(SD = 12.37)*	*(SD = 10.99)*	*(SD = 8.89)*	*(SD = 13.29)*
Male	$N = 41$	$N = 37$	$N = 24$	$N = 15$	$N = 27$
Female	$N = 0$	$N = 0$	$N = 6$	$N = 8$	$N = 20$
IQ*	$M = 89.85^{a,b}$	$M = 89.85^{a,b}$	$M = 88.83^{a,b,c}$	$M = 95.43^{a,c,d}$	$M = 100.18^{d}$
	(SD = 11.25)	*(SD = 11.25)*	*(SD = 11.70)*	*(SD = 10.72)*	*(SD = 11.89)*
Weekly alcohol use	$M = 9.63$	$M = 4.49$	$M = 5.23$	$M = 1.65$	$M = 5.91$
	(SD = 23.53)	*(SD = 7.55)*	*(SD = 10.36)*	*(SD = 2.48)*	*(SD = 5.74)*
Weekly cannabis use	$M = 6.24$	$M = .54$	$M = 4.10$	$M = .74$	$M = .02$
	(SD = 22.43)	*(SD = 2.42)*	*(SD = 10.74)*	*(SD = 2.07)*	*(SD = .15)*
AQ total	$M = 95.85^{a}$	$M = 80.86^{b}$	$M = 99.90^{a}$	$M = 77.35^{b}$	$M = 57.83^{c}$
	(SD = 13.84)	*(SD = 17.05)*	*(SD = 16.58)*	*(SD = 12.46)*	*(SD = 10.38)*
HIT total	$M = 25.73^{a}$	$M = 20.76^{b,c}$	$M = 23.73^{a,b}$	$M = 17.10^{c,d}$	$M = 17.10^{d}$
	(SD = 6.30)	*(SD = 5.74)*	*(SD = 7.65)*	*(SD = 4.47)*	*(SD = 4.47)*
Proactive aggression (RPQ)	$M = 5.80^{a}$	$M = 2.16^{b,c}$	$M = 3.83^{b}$	$M = 1.65^{b,c}$	$M = .72^{b,c}$
	(SD = 3.85)	*(SD = 2.59)*	*(SD = 3.53)*	*(SD = 1.95)*	*(SD = 1.27)*
Reactive aggression (RPQ)	$M = 14.07^{a}$	$M = 9.81^{b}$	$M = 13.20^{a}$	$M = 8.39^{b}$	$M = 4.59^{c}$
	(SD = 4.31)	*(SD = 3.99)*	*(SD = 4.32)*	*(SD = 4.19)*	*(SD = 2.93)*
SDAS total	$M = 15.10^{a}$	$M = 12.57^{a}$	$M = 18.87^{b}$	$M = 8.30^{c}$	$M = 2.85^{d}$
	(SD = 5.90)	*(SD = 5.32)*	*(SD = 6.83)*	*(SD = 4.26)*	*(SD = 2.36)*

*As measured by using the Dutch Adult Reading Test (Schmand, Bakker, Saan, & Louman, 1991).
Superscripts of mean values indicate significant group differences: Groups with different superscripts differ from each other at least $P < .05$.

5 = extremely like me). The Dutch translation has adequate psychometric properties (Cronbach's Alpha = .86; Morren & Meesters, 2002). In the present study the internal consistency has also proven to be good (Cronbach's Alpha = .93).

The *How i think questionnaire* (HIT; Barriga & Gibbs, 1996) is a 54-item self-report questionnaire to assess self-serving cognitive distortions. The items are divided into four cognitive distortion subscales (self-centered, blaming others, minimizing/labeling, assuming the worst) and four behavioral referent categories (physical aggression, opposition-defiance, lying, stealing). Items have to be answered on a 6-point Likert scale (1 = totally agree to 6 = totally disagree). The Dutch translation has proven reliability (Cronbach's Alpha ranged from .90 to .94) and validity (all $r < .20$; Nas, Brugman, & Koops, 2008). In the current study the internal consistency has also proven to be good (Cronbach's Alpha = .94).

Hostile Interpretation Bias Task

The *Hostile interpretation bias task* (HIBT) was developed to assess a HIB. Photos of faces with emotional affect (angry, fear, disgust, happy) of four male and four female models were selected from the Radboud Faces Database (Langner et al., 2010). Each affective picture was morphed (using WinMorph 3.01) five times with the neutral image of the same individual, creating 20%, 40%, 60%, 80%, and 100% emotion intensity, respectively. The neutral expression was in all models displayed with mouth closed whereas the emotional pictures were displayed with mouth open. This difference in mouth opening resulted in pictures showing ambiguous expressions. An example is displayed in Figure 1.

The task consisted of a practice block and two experimental blocks. The practice block consisted of 16

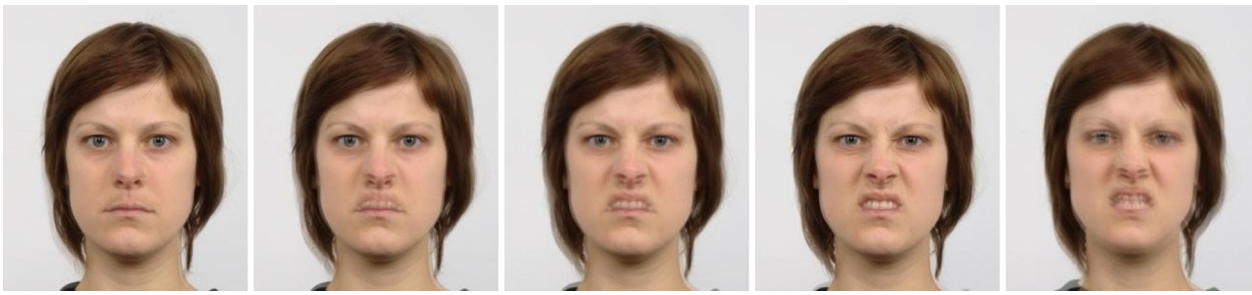

Fig. 1. Example of morphed emotional facial expressions.

trials (8 models × 2 emotions). Only pictures with happy and angry affect and of 100% intensity were used to familiarize participants with the task. Each experimental block consisted of 168 trials (8 models × 4 emotions × 5 intensity levels + 8 neutral images). The order of the pictures was randomized and equal in both blocks and equal to every participant. Participants were instructed to indicate whether the picture looked hostile or not. In case they thought they saw a hostile picture, they were asked to press the Z-key, otherwise the M-key (on a qwerty keyboard). They had to respond as quickly as possible. The picture, size 8.5 × 10.5 cm, was presented for four seconds, in the center of the computer screen, against a black background. The pictures remained on the screen until a response was given or until 4 sec had passed. The time period of 4 sec was chosen to let participants both observe the picture and respond based on their impulses. After a pretrial pause of 1 sec, a new picture was displayed immediately. Labels were displayed in the left (Yes, hostile) and right (No, not hostile) bottom corner of the screen in white Arial font, size 30. Responses given by pressing the Z-key, indicating that the participant saw a hostile picture, were defined as "hostile" responses. If a response was not given within 4 sec, the words "Too late" appeared on the screen in red. A hostile interpretation bias was defined as the percentage of "hostile" responses to the emotional pictures. The hostile responses were dummy coded (0 = no, not hostile, 1 = yes, hostile), and the mean was calculated: this immediately revealed the percentage of the pictures that were interpreted as hostile. Trials without a response (due to late responding) were not taken into account. In total, it took participants approximately 10 min to complete the HIBT.

Furthermore, the HIBT was used in a small pilot study ($N = 22$) in which forensic psychiatric inpatients were assessed. These patients were all diagnosed with antisocial personality disorder and did not participate in the current study. In this pilot study, the HIBT was administered twice with a 1-week interval. The correlations of the two measurements for angry, fear, happy, and disgust were $r = .774$, $P < .001$; $r = .904$, $P < .001$; $r = .295$, $P = .182$; $r = .908$, $P < .001$, respectively. These results suggest good test-retest reliability, except for faces displaying happy affect.

RESULTS

Group Differences on HIB

The mean percentage of "hostile" responses per intensity and emotion is displayed in Figure 2. To analyze whether FPOs (ASPD, IED, BPD) show a different pattern of hostile interpretation than nFPOs-BPD and HCs, first an overall omnibus test was conducted with a 4 (emotion: anger, happy, fear, disgust) × 5 (intensity: 20, 40, 60, 80, 100%) × 5 (group: ASPD vs. IED vs. BPD vs. nFPOs-BPD vs. HCs) repeated-measures design. Due to violation of sphericity, Greenhouse-Geisser correction was used. Bonferroni correction was used to control for multiple comparisons. Significant main effects of *emotion*, *intensity*, and *group* were found, $F(2.43, 417.11) = 415.41$, $P < .001$, $eta^2 = .707$; $F(1.67, 287.57) = 216.80$, $P < .001$, $eta^2 = .558$; and $F(4, 172) = 5.53$, $P < .001$, $eta^2 = .114$, respectively. The *intensity* × *group* and *emotion* × *intensity* × *group* interactions were not significant, $F(6.69, 287.57) = 1.284$, $P = .260$, $eta^2 = .029$; and $F(30.37, 1305.89) = 1.41$, $P = .072$, $eta^2 = .032$, respectively. The *emotion* × *group* and *emotion* × *intensity* interactions did reach significance, $F(9.70, 417.11) = 1.284$, $P = .046$, $eta^2 = .042$; and $F(7.59, 1305.89) = 220.42$, $P < .001$, $eta^2 = .562$, respectively. These results indicate that the interpretation of hostility differed for each emotion and intensity. Faces displaying anger and disgust were interpreted as hostile more often than expressions with fear or happy affect (mean percentage "hostile" responses; anger: 52.89%; disgust: 45.57%; fear: 21.58%, happy: 7.51%). Moreover, happy faces were the least interpreted as hostile. Regarding intensity, Figure 2 illustrates that for angry, disgust, and fearful expressions, "hostile" interpretations increased with increasing intensity of the emotion, whereas the opposite was true for expressions of happiness. The results also suggest that FPOs (ASPD, IED, BPD), nFPOs-BPD, and HCs differed in their overall level of perceived hostility and in their patterns of hostile interpretations of emotional facial expressions. There was a significant difference between the FPOs (ASPD, IED, BPD), nFPOs-BPD and HC for IQ and gender (see Table I). To analyze whether IQ and gender functioned as a confounding variables, the analysis described above was conducted again, now with IQ as a covariate and gender as an additional between-subjects factor. Results showed that the effect sizes did not differ essentially when IQ and gender were taken into account. It was also revealed that IQ was not associated with the frequency of "hostile" responses, whereas gender was, $F(1, 166) = 2.29$, $P = .131$, $eta^2 = .014$; $F(1, 166) = 5.17$, $P = .024$, $eta^2 = .03$, respectively. Therefore, in subsequent analyses only gender was included as an additional between-subjects factor.

Group Differences on HIB Per Emotion

As the previous analysis revealed an interaction effect between emotion and group, four different 5 (intensity: 20, 40, 60, 80, 100%) × 5 (group: ASPD vs. IED vs. BPD vs. nFPOs-BPD vs. HCs) × 2 gender (male, female)

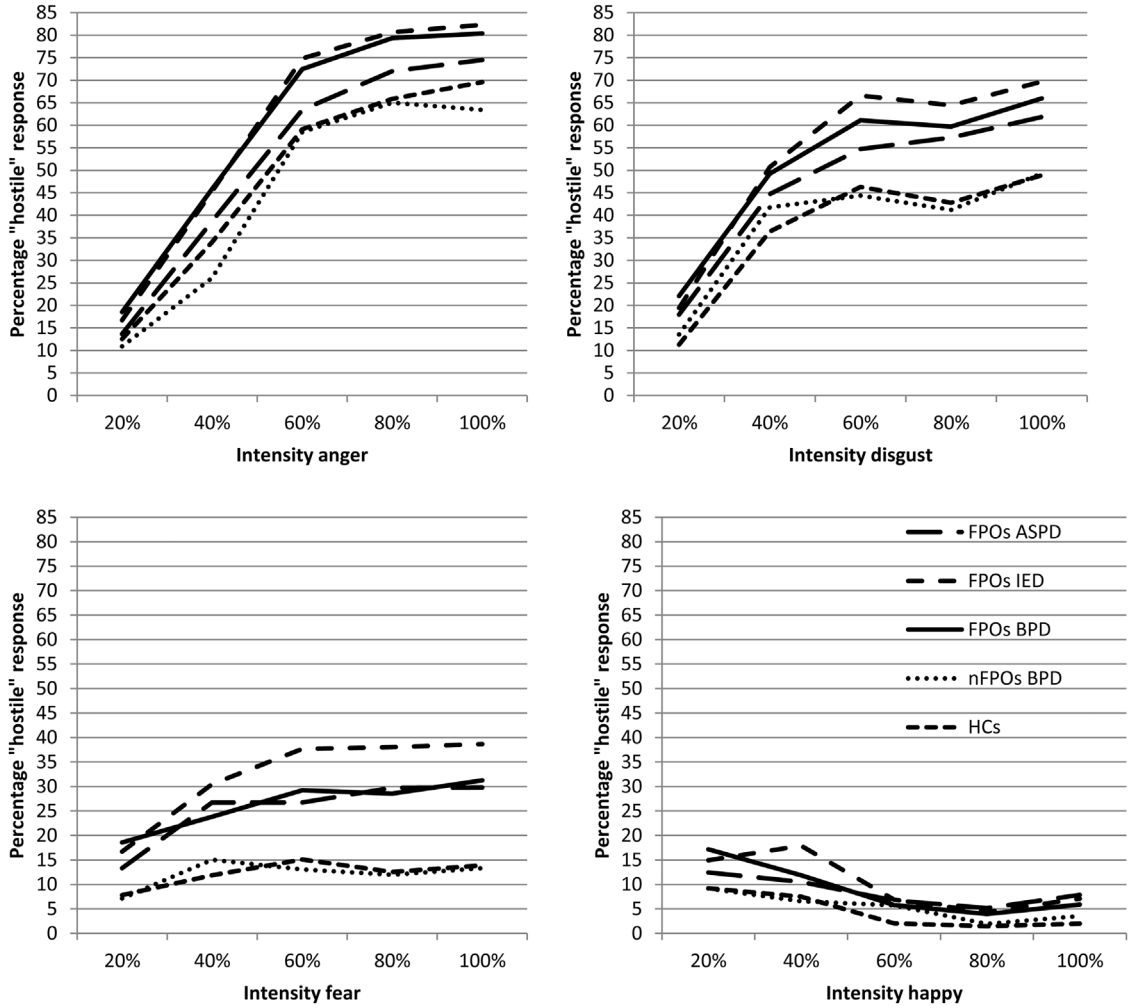

Fig. 2. Percentage "hostile" response per emotion condition.

repeated-measures ANOVAs were conducted, one for each emotion, to analyze more specifically on which emotions FPOs (ASPD, IED, BPD), nFPOs-BPD, and HCs differed. Due to violation of sphericity, Greenhouse-Geisser correction was used. To control for multiple comparisons, Bonferroni correction was used in all subsequent analyses. There were significant main effects of intensity and group for anger, $F(2.26, 382.37) = 275.17$, $P < .001$, $eta^2 = .62$; and $F(4, 169) = 3.34$, $P = .012$, $eta^2 = .073$, respectively. The *intensity* × *group* interaction was non-significant: $F(9.05, 382.37) = .71$, $P = .700$, $eta^2 = .017$. Gender was not a significant covariate: $F(1, 169) = 1.88$, $P = .173$, $eta^2 = .011$.

The same was found for disgust: Significant main effects of *intensity* and *group* were observed, $F(2.28, 385.87) = 113.38$, $P < .001$, $eta^2 = .402$; and $F(4, 169) = 4.40$, $P = .002$, $eta^2 = .094$, respectively. The *intensity* × *group* interaction was not significant, $F(9.13, 385.87) = 1.37$, $P = .197$, $eta^2 = .031$. Gender did function as a significant covariate, $F(1, 169) = 4.09$, $P = .045$, $eta^2 = .024$.

There were significant main effects of *intensity* and *group* for fear, $F(1.92, 324.55) = 16.51$, $P < .001$, $eta^2 = .089$; and $F(4, 169) = 7.15$, $P < .001$, $eta^2 = .145$, respectively. The *intensity* × *group* interaction also reached significance: $F(7.68, 324.55) = 2.23$, $P = .027$, $eta^2 = .05$. Gender was a significant covariate: $F(1, 169) = 5.75$, $P = .018$, $eta^2 = .033$.

For happiness, significant main effects of *intensity* and *group* were found again: $F(2.55, 431.71) = 26.75$, $P < .001$, $eta^2 = .137$; and $F(4, 169) = 3.58$, $P = .008$, $eta^2 = .078$, respectively. The *intensity* x *group* interaction also reached significance: $F(10.22, 431.71) = 2.17$, $P = .018$, $eta^2 = .049$. Gender did not function as a significant covariate: $F(1, 169) = 2.07$, $P = .152$, $eta^2 = .012$.

Taken together, these results indicate that angry, disgusted, and fearful pictures of higher intensity were

more often interpreted as hostile whereas high intensities of happiness were interpreted as less hostile than low intensities. Moreover, gender only functioned as a significant moderator for the interpretation of fearful and disgusted faces. However, even here the main effect of group remained significant. In addition, group-wise comparisons indicated that FPOs (ASPD, BPD, IED) differed from HCs on all four emotional expressions (anger: $P = .046$; happy: $P = .006$; fear: $P < .001$; disgust: $P = .004$). FPOs differed from nFPOs-BPD solely on faces displaying anger or fear (anger: $P = .034$; fear: $P = .007$). The nFPOs-BPD did not differ from HCs on any of the emotions. In addition, the three groups of FPOs did not differ from each other on any of the emotions.

Correlations Between HIB and Questionnaires

To explore the associations between HIB and type and severity of aggression, trait aggression, and cognitive distortions, correlational analyses were performed on the full sample. The two subscales of the RPQ and the total scores of the other questionnaires were of interest in the current analyses. The correlation between the raw mean reactive and raw mean proactive score of the RPQ was high ($r = .75$, CI = .69−.80, $P < .001$). Therefore, the residualized measures of both subscales were created in order to assess the correlates of reactive and proactive aggression independently of one another, as suggested by Raine et al. (2006) and Cima et al. (2013). Reactive aggression was regressed on proactive scores and Pearson standardized residuals (with a mean of 0 and SD of 1) were saved to index residualized proactive aggression, while the standardized residuals of proactive aggression on reactive aggression were saved to index residualized reactive aggression.

Correlations were computed using the percentage of "hostile" responses per emotion condition, averaging across intensities. To determine confidence intervals and to test the significance of the correlations, a bootstrapping (1,000 samples) procedure was used. Bootstrapping is based on random sampling with replacement. It simulates the population distribution of the correlation and to provides confidence intervals for the correlation coefficients (Sideridis & Simos, 2010). When using this approach a more accurate estimate of the associations is provided than estimates produced by a single sample (Hesterberg, Monaghan, Moore, Clipson, & Epstein, 2003).

Correlations are displayed in Table II. A HIB regarding angry faces was associated with the total AQ score, total HIT score, and total SDAS score. "Hostile" responses to happy faces were related to total HIT score, and total SDAS score. A HIB regarding fearful faces correlated with total AQ score, total HIT score, reactive aggression (RPQ), and total SDAS score. There was a significant association between "hostile" responses to disgusted faces and total AQ score, total HIT score, reactive aggression (RPQ), and total SDAS score.

Group Differences on Questionnaires

As the correlational analysis revealed that there was an association between a HIB regarding facial expressions and type and severity of aggression, and cognitive distortions, a MANOVA was conducted to investigate whether FPOs differed in these characteristics from nFPOs-BPD and HCs. Means are displayed in Table I. Bonferroni correction was used to control for multiple comparisons. A significant multivariate effect of group was found, Wilks' Lambda = .280, $F(20, 554.83) = 12.96$, $P < .001$, $eta^2 = .272$. Separate univariate ANOVAs showed significant group differences for total AQ score, total HIT score, total ISS score, proactive aggression (RPQ), reactive aggression (RPQ), and total SDAS score: $F(4, 175) = 55.73$, $P < .001$, $eta^2 = 566$; $F(4, 175) = 25.59$, $P < .001$, $eta^2 = .375$; $F(4, 175) = 20.34$, $P < .001$, $eta^2 = .322$; $F(4, 175) = 38.85$, $P < .001$, $eta^2 = .476$; and $F(4, 175) = 57.95$, $P < .001$, $eta^2 = .574$, respectively. These results showed that FPOs with ASPD and BPD had a higher disposition to act aggressively, they had higher levels of cognitive distortions, proactive aggression, and reactive aggression, and they showed aggressive behavior more frequently than nFPOs-BPD and HCs. The FPOs with IED displayed similar scores on all questionnaires as the nFPOs-BPD, except for the SDAS. The nFPOs-BPD showed more aggressive behavior, except reactive aggression, than HCs (see Table I).

DISCUSSION

The current study found a highly generalized HIB to facial expressions among FPOs diagnosed with ASPD, BPD, or IED, rather than nFPOs- BPD and HCs. The present findings are in agreement with those of Schonenberg and Jusyte (2014), who found that antisocial violent offenders interpreted ambiguous facial cues as hostile. Previous studies suggested that aggressive individuals tend to attribute hostility to others in socially ambiguous situations (Bailey & Ostrov, 2007; Crick, 1995; Crick et al., 2002; De Castro et al., 2002; Dodge, 2006; Lobbestael et al., 2013). Together with the study by Schonenberg and Jusyte (2014), the current research provides support for the idea that a hostile attribution bias exists not only regarding ambiguous social situations, but also with respect to ambiguous and less ambiguous facial expressions.

Overall, a significant effect of intensity was found; as the intensity of facial expressions displaying anger,

TABLE II. Correlations Between HIB and Questionnaires

Questionnaire	HIB angry faces	HIB happy faces	HIB fearful faces	HIB disgusted faces
AQ total	$r = .219$, *CI = .072–.365,* $P = .004$	$r = .203$, *CI = .042–.354,* $P = .007$	$r = .182$, *CI = .031–.322,* $P = .016$	$r = .186$, *CI = .040–.321,* $P = .014$
HIT total	$r = .176$, CI = .008–.345, $P = .020$	$r = .286$, CI = .145–.432, $P < .001$	$r = .203$, *CI = .069–.335,* $P = .007$	$r = .172$, *CI = .023–.328,* $P = .022$
Residual proactive aggression (RPQ)	n.s	n.s	n.s	n.s
Residual reactive aggression (RPQ)	n.s	n.s	$r = .162$, *CI = .011–.299,* *P =.032*	$r = .212$, *CI = .077–.344* $P = .005$
SDAS total	$r = .193$, CI = .064–.323, $P = .010$	$r = .263$, *CI = .090–.416,* $P < .001$	$r = .256$, *CI = .104–.389,* $P < .001$	$r = .222$, *CI = .084–.354,* $P = .003$

n.s., non significant.

disgust or fear increased, expressions were more often interpreted as hostile. These results indicate that the interpretation bias of hostility is not limited to angry expressions, as may have been suggested by previous studies. This intensity effect of disgusted faces was also found among HCs. Disgust and contempt look very similar. However, disgust is often displayed regarding objects, whereas contempt relates more to people. Perhaps disgust is misinterpreted as contempt and therefore experienced as hostile. Additionally, it is also conceivable that anger and disgust are confused with each other based on the similarities between these expressions (Wieser & Brosch, 2012). The influence of contextual cues on the processing of facial expressions is considered to be significant (Wieser & Brosch, 2012). As the current task only presented facial expressions without any contextual cues, it is possible that faces displaying anger and disgust were evaluated alike. The intensity effect of fear, on the other hand, may be explained in terms of recognition deficits, as antisocial behavior was associated with deficits in the recognition of fearful facial affect (Marsh & Blair, 2008). It might be suggested that FPOs process fearful expressions incorrectly and therefore display biased interpretations. The intensity effect was reversed for happy faces; as the intensity increased, the faces were less often interpreted as hostile. This is plausible as happy faces are probably experienced as more positive and therefore as less hostile.

In line with previous research, which suggested that aggressive individuals attribute hostility to others in socially ambiguous situations, it was predicted that FPOs only differed in interpretation of hostility for more ambiguous facial expressions (40%, 60%, 80%) and not in the least ambiguous pictures with 20% or 100% intensity of the facial expression. The current results, however, revealed that FPOs exhibited a comparable pattern of HIB across all intensity levels. This indicates that FPOs interpret not only ambiguous expressions, but emotional facial expressions in general as hostile

It was revealed that all FPOs displayed more "hostile" responses to angry, disgusted, fearful, and happy facial expressions than HCs did. The FPOs interpreted faces displaying anger or disgust more often as hostile than nFPOs-BPD. No differences were found between responses on the HIBT between nFPOs-BPD and HCs. Based on previous literature, which suggested that ASPD, BPD, and IED all are likely to exhibit cognitive and interpretation biases, it was predicted that not only the FPOs would show a HIB, but also the nFPOs- BPD. Although the nFPOs-BPD consisted of a group of psychiatric patients receiving intensive treatment, they performed surprisingly similar to the HCs on the HIBT. As both the FPOs and nFPOs- BPD were assessed before the start of their treatment, possible treatment effects on this hostility bias were ruled out. The present findings strongly suggest that the FPOs generally tend to interpret various emotions, at every intensity level, more frequently as hostile than nFPOs-BPD or HCs do. The FPOs exhibited much higher levels of aggressive behavior, the findings lend support to the notion that a HIB regarding facial expressions is a characteristic of pathological aggression, typically shown in forensic settings, regardless of a psychiatric diagnosis of ASPD, BPD, or IED.

Previous studies have shown that there was an association between hostility biases and reactive aggressive behavior (Bailey & Ostrov, 2007; Chen et al., 2012; Crick, 1995; Crick et al., 2002; De Castro et al., 2002; Dodge, 2006; Lobbestael et al., 2013). In accordance with these studies, the current study revealed an association with reactive aggression. The correlation, however, only occurred for faces displaying fear and disgust. No link was observed with proactive aggression.

Clear associations were found between the overall HIB and severity of aggressive behavior and for the overall trait of aggression. This might be explained by the high levels of aggressive behavior, in general, displayed by the FPOs; not only higher levels of reactive and proactive aggression and trait aggression were displayed by FPOs with ASPD and BPD than nFPOs-BPD and HCs, but also more severe aggressive behavior was shown by all three groups of FPOs, that is, they showed more pathological forms of aggressive behavior. The current results underline the idea that individuals with aggression regulation problems, in real life, have an increased disposition to respond aggressively to ambiguous situations as well as to ambiguous and less ambiguous facial expressions.

A HIB was also found to be associated with another clinically relevant characteristic: self-serving cognitive distortions. It is well known that cognitive distortions are related to violent and aggressive behavior (Chereji, Pintea, & David, 2012). However, the present study is one of the first to show a link between conscious self-serving cognitive distortions and a more implicit HIB regarding emotional facial expressions. The current result underlined the external validity of the HIBT. In addition, this finding might suggest the notion of a general predisposition in FPOs to display cognitive biases, whether it is conscious or more automatically. This association needs to be studied in further detail. It underlines the necessity to not only alter conscious cognitive distortions but to also focus on automatic biases in order to reduce aggressive behavior.

The current findings may have some implications with respect to forensic clinical practice. The HIB might be the result of a failure to learn how to make a benign interpretation, or due to deficits in the evaluation and appraisal of situations, as is suggested by the Developmental Model of Hostile Attribution Style, the Social Information-Processing Model, and the General Aggression Model (Crick & Dodge, 1994; DeWall, Anderson, & Bushman, 2011; Dodge, 2006). The current results emphasize the need to determine the presence of this bias in order to reduce it in all variants of interventions. Besides developing skills to control anger, FPOs need to become aware of their automatic tendency to interpret facial expressions as hostile and to learn how to interpret facial expressions differently and more accurately.

The present study and has a number of limitations which merit further comment. First, except for the HIBT, all measurements consisted of self-report. It is questionable whether a population of FPOs is fully able to reflect on their own behavior and whether they are willing to answer genuinely. Second, a HIB may be closely related to deficits in facial emotion recognition. Regarding prospective research, it would be interesting to include an emotion recognition task to explore its associations with a bias in the interpretation of facial affect. Third, it would be of great importance to include another task with a different response format. Hence, it would be possible to exclude any response bias which may have occurred in the current study, as the FPOs tended to show more "yes" responses. Fourth, it may be possible that FPOs are better at detecting potential threat and might, therefore, interpret angry and disgusted faces correctly as hostile instead of displaying a bias. More research, however, is needed to explore this possibility. Fifth, in the current task only facial expressions displaying anger, fear, disgust, and happiness were selected. For future research, it will be necessary to include other emotional facial expressions to investigate the generalizability of this hostility bias. Sixth, the non-forensic group consisted solely of patients with BPD, and the sample size of this group was rather small. The gender distribution varied in the different groups, and women were underrepresented overall (19.2%). The results, therefore, are most applicable to male individuals. Future research using larger samples should replicate the current findings.

Notwithstanding the limitations, this study revealed the existence of a generalized HIB of emotional facial expressions which is highly likely to be a characteristic of pathological aggressive behavior displayed by FPOs. Future research could investigate the causality of the association between pathological aggressive behavior and HIB. At present there are multiple types of hostility biases that have been identified by very different paradigms of which the hostile attribution bias is the most widely researched version. Our HIB and also the hostile expectation bias found by Bushman and Anderson (2002), are other examples of hostility biases. All these biases have been repeatedly found to be associated with higher levels of aggressive, violent and antisocial behavior. To date, however, there is no evidence on whether these biases are distinct phenomena or whether they are manifestations of the same underlying construct, this will be an important avenue for future research. Subsequently, more knowledge is needed on how to reduce these biases successfully. Previous research showed that it was possible to modify biases in emotion recognition and to improve facial affect recognition after a brief computerized training (Penton-Voak et al., 2013; Schönenberg et al., 2014). It would be of great interest to assess whether these techniques are able to reduce a general HIB. In the long run, interventions that alter the HIB might even help to reduce the recurrence of aggression.

ACKNOWLEDGMENTS

This research was funded by NWO/National Initiative Brain & Cognition (NIHC, grant NWO-056-24-011). The authors would like to thank Ricardo Pengel of Radboud UMC and the staff in the outpatient clinics of Kairos, Pompestichting at Nijmegen, The Netherlands, for their support in recruiting participants.

REFERENCES

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders (4th ed., Text revision, DSM– IV–TR)*. Arlington, VA: APA.

Arntz, A., Weertman, A., & Salet, S. (2011). Interpretation bias in Cluster-C and borderline personality disorders. *Behaviour Research and Therapy*, *49*, 472–481. doi: doi.org/10.1016/j.brat.2011.05.002

Baer, R. A., Peters, J. R., Eisenlohr-Moul, T. A., Geiger, P. J., & Sauer, S. E. (2012). Emotion-related cognitive processes in borderline personality disorder: A review of the empirical literature. *Clinical Psychology Review*, *32*, 359–369. doi: 10.1016/j.cpr.2012.03.002

Bailey, C. A., & Ostrov, J. M. (2007). Differentiating forms and functions of aggression in emerging adults: Associations with hostile attribution biases and normative beliefs. *Journal of Youth and Adolescence*, *37*, 713–722. doi: 10.1007/s10964-007-9211-5

Barriga, A. Q., & Gibbs, J. C. (1996). Measuring cognitive distortion in antisocial youth: Development and preliminary validation of the "How I Think" questionnaire. *Aggressive Behavior*, *22*, 333–343. doi:org/10.1002/(SICI)1098-2337(1996)22:5%3C333::AID-AB2%3E3.0.CO;2-K

Burt, S. A., Mikolajewski, A. J., & Larson, C. L. (2009). Do aggression and rule-breaking have different interpersonal correlates? A study of antisocial behavior subtypes, negative affect, and hostile perceptions of others. *Aggressive Behavior*, *35*, 453–461. doi: 10.1002/ab.20324

Bushman, B. J., & Anderson, C. A. (2002). Violent video games and hostile expectations: A test of the general aggression model. *Personality and Social Psychology Bulletin*, *28*, 1679–1686. doi: org/10.1177/014616702237649

Buss, A. H., & Perry, M. (1992). The aggression questionnaire. *Journal of Personality and Social Psychology*, *63*, 452. doi: 10.1037/ 0022-3514.63.3.452

Chen, P., Coccaro, E. F., & Jacobson, K. C. (2012). Hostile attributional bias, negative emotional responding, and aggression in adults: Moderating effects of gender and impulsivity. *Aggressive Behavior*, *38*, 47–63. doi: 10.1002/ab.21407

Chereji, S. V., Pintea, S., & David, D. (2012). The relationship of anger and cognitive distortions with violence in violent offenders' population: A meta-analytic review. *The European Journal of Psychology Applied to Legal Context*, *4*, 59–77.

Cima, M., Raine, A., Meesters, C., & Popma, A. (2013). Validation of the dutch reactive proactive questionnaire (RPQ): Differential correlates of reactive and proactive aggression from childhood to adulthood. *Aggressive Behavior*, *39*, 99–113. doi: 10.1002/ab.21458

Coccaro, E. F., Kavoussi, R. J., Berman, M. E., & Lish, J. D. (1998). Intermittent explosive disorder-revised: Development, reliability, and validity of research criteria. *Comprehensive Psychiatry*, *39*, 368–376. doi: 10.1016/S0010-440X(98)90050-5

Coccaro, E. F., Solis, O., Fanning, J., & Lee, R. (2015). Emotional intelligence and impulsive aggression in Intermittent Explosive Disorder. *Journal of Psychiatric Research*, *61*, 135–140. doi: org/10.1016/j.jpsychires.2014.11.004

Crick, N. R. (1995). Relational aggression: The role of intent attributions, feelings of distress, and provocation type. *Developmental Psychopathology*, *7*, 313–313. doi: 10.1017/S0954579400006520

Crick, N. R., & Dodge, K. A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychological Bulletin*, *115*, 74. doi: 10.1037/0033-2909. 115.1.74

Crick, N. R., & Dodge, K. A. (1996). Social information-processing mechanisms in reactive and proactive aggression. *Child Development*, *67*, 993–1002. doi: 10.1111/j.1467-8624.1996.tb01778.x

Crick, N. R., Grotpeter, J. K., & Bigbee, M. A. (2002). Relationally and physically aggressive children's intent attributions and feelings of distress for relational and instrumental peer provocations. *Child Development*, *73*, 1134–1142. doi: 10.1111/ 1467-8624.00462

Daros, A. R., Zakzanis, K. K., & Ruocco, A. C. (2013). Facial emotion recognition in borderline personality disorder. *Psychological Medicine*, *43*, 1953–1963. doi: 10.1017/S0033291712002607

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 & Biobehavioral Reviews*, *36*, 2288–2304.

De Castro, B. O., Veerman, J. W., Koops, W., Bosch, J. D., & Monshouwer, H. J. (2002). Hostile attribution of intent and aggressive behavior: A meta-analysis. *Child Development*, *73*, 916–934. doi: 10.1111/ 1467-8624.00447

DeWall, C. N., Anderson, C. A., & Bushman, B. J. (2011). The general aggression model: Theoretical extensions to violence. *Psychology of Violence*, *1*, 245. doi: 10.1037/a0023842

Dodge, K. A. (2006). Translational science in action: Hostile attributional style and the development of aggressive behavior problems. *Developmental Psychopathology*, *18*, 791–814. doi: 10.10170S0954579406060391

Dodge, K. A., Bates, J. E., & Pettit, G. S. (1990). Mechanisms in the cycle of violence. *Science*, *250*, 1678–1683. doi: 10.1126/science.2270481

Dolan, M., & Fullam, R. (2006). Face affect recognition deficits in personality-disordered offenders: Association with psychopathy. *Psychological Medicine*, *36*, 1563–1569. doi: 10.1017/S0033291706008634

Eckhardt, C., Norlander, B., & Deffenbacher, J. (2004). The assessment of anger and hostility: A critical review. *Aggression and Violent Behavior*, *9*, 17–43. doi: 10.1016/s1359-1789(02)00116-7

Epstein, S., & Taylor, S. P. (1967). Instigation to aggression as a function of degree of defeat and perceived aggressive intent of the opponent1. *Journal of Personality*, *35*, 265–289. doi: 10.1111/j. 1467-6494.1967.tb01428.x

Fertuck, E. A., Grinband, J., & Stanley, B. (2013). Facial trust appraisal negatively biased in borderline personality disorder. *Psychiatry Research*, *207*, 195–202. doi: 10.1016/j.psychres.2013.01.004

Helfritz-Sinville, L. E., & Stanford, M. S. (2014). Hostile attribution bias in impulsive and premeditated aggression. *Personality and Individual Differences*, *56*, 45–50. doi: 10.1016/j.paid.2013.08.017

Hesterberg, T., Monaghan, S., Moore, D., Clipson, A., & Epstein, R. (2003). *Bootstrap methods and Permutation Tests: Companion chapter 18 to The Practice of Business Statistics*. New York: W.H. Freemann and Company.

Hoaken, P. N., Allaby, D. B., & Earle, J. (2007). Executive cognitive functioning and the recognition of facial expressions of emotion in incarcerated violent offenders, non-violent offenders, and controls. *Aggressive Behavior*, *33*, 412–421. doi: 10.1002/ab.20194

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the radboud faces database. *Cognition and Emotion*, *24*, 1377–1388. doi: 10.1080/02699930903485076

Linehan, M., & Read, K. (2013). *Instructor's manual for dialectical behavioral therapy*. Mill Valley, CA: Psychotherapy. net.

Lobbestael, J., Cima, M., & Arntz, A. (2013). The relationship between adult reactive and proactive aggression, hostile interpretation bias, and antisocial personality disorder. *Journal of Personality Disorders*, *27*, 53–66. doi: 10.1521/pedi.2013.27.1.53

Lobbestael, J., & McNally, R. J. (2015). An empirical test of rejection- and anger-Related interpretation bias in borderline personality disorder. *Journal of Personality Disorders*, 1–13. doi: org/10.1521/pedi_2015_29_194

Marsh, A. A., & Blair, R. (2008). Deficits in facial affect recognition among antisocial populations: A meta-analysis. *Neuroscience & Biobehavioral Reviews*, *32*, 454–465. doi: 10.1016/j.neubiorev.2007.08.003

Mayer, J. D., & Salovey, P. (1993). The intelligence of emotional intelligence. *Intelligence*, *17*, 433–442. doi: 10.1016/0160-2896(93)90010-3

Moriya, J., Tanno, Y., & Sugiura, Y. (2012). Repeated short presentations of morphed facial expressions change recognition and evaluation of facial expressions. *Psychological Research*, *77*, 698–707. doi: 10.1007/s00426-012-0463-7

Morren, M., & Meesters, C. (2002). Validation of the Dutch version of the Aggression Questionnaire in adolescent male offenders. *Aggressive Behavior*, *28*, 87–96. doi: 10.1002/ab.90010

Nas, C. N., Brugman, D., & Koops, W. (2008). Measuring self-serving cognitive distortions with the "How I think" questionnaire. *European Journal of Psychological Assessment*, *24*, 181–189.

Nasby, W., Hayden, B., & DePaulo, B. M. (1980). Attributional bias among aggressive boys to interpret unambiguous social stimuli as displays of hostility. *Journal of Abnormal Psychology*, *89*, 459. doi: 10.1037/0021-843X.89.3.459

Penton-Voak, I. S., Thomas, J., Gage, S. H., McMurran, M., McDonald, S., & Munafò, M. R. (2013). Increasing recognition of happiness in ambiguous facial expressions reduces anger and aggressive behavior. *Psychological Science*, *24*, 688–697. doi: 10.1177/0956797612459657

Raine, A., Dodge, K., Loeber, R., Gatzke-Kopp, L., Lynam, D., Reynolds, C., & Liu, J. H. (2006). The reactive-proactive aggression questionnaire: Differential correlates of reactive and proactive aggression in adolescent boys. *Aggressive Behavior*, *32*, 159–171. doi: 10.1002/ab.20115

Schmand, B. A., Bakker, D., Saan, R. J., & Louman, J. (1991). De Nederlandse Leestest voor Volwassenen: Een maat voor het premorbide intelligentieniveau. *Tijdschrift voor Gerontologie en Geriatrie*, *22*, 15–19.

Schönenberg, M., Christian, S., Gaußer, A.-K., Mayer, S., Hautzinger, M., & Jusyte, A. (2014). Addressing perceptual insensitivity to facial affect in violent offenders: First evidence for the efficacy of a novel implicit training approach. *Psychological Medicine*, *44*, 1043–1052.

Schonenberg, M., & Jusyte, A. (2014). Investigation of the hostile attribution bias toward ambiguous facial cues in antisocial violent offenders. *European Archives of Psychiatry and Clinical Neurosciences*, *264*, 61–69. doi: 10.1007/s00406-013-0440-1

Sheehan, D., Lecrubier, Y., Sheehan, K., Amorim, P., Janavs, J., & Weiler, E. (1998). The Mini-International Neuropsychiatric Interview (MINI): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry*, *59*, 22–33. doi: 10.1016/s0924-9338(97)83296-8

Sideridis, G. D., & Simos, P. (2010). What is the actual correlation between expressive and receptive measures of vocabulary? Approximating the sampling distribution of the correlation coefficient using the bootstrapping method. *International Journal of Educational and Psychological Assessment*, *5*, 117–133. doi: 10.1044/1058-0360(2009/08-0035)

Siever, L. J. (2008). Neurobiology of aggression and violence. *American Journal of Psychiatry*, *165*, 429–442. doi: org/10.1176/appi.ajp.2008.07111774

Van Vliet, I., & De Beurs, E. (2007). Het mini internationaal neuropsychiatrisch interview (MINI). een kort gestructureerd diagnostisch psychiatrisch interview voor DSM-IV en ICD-10-stoornissen [The mini international neuropsychiatric interview (MINI). a short structured diagnostic psychiatric interview for DSM-IV and ICD-10 disorders]. *Tijdschrift voor Psychiatrie*, *49*, 393–397.

Weertman, A., Arntz, A., & Kerkhofs, M. (2000). *SCID-II: Gestructureerd klinisch interview voor DSM-IV as II persoonlijkheidsstoornissen*. Lisse: Swets & Zeitlinger.

Wieser, M. J., & Brosch, T. (2012). Faces in context: A review and systematization of contextual influences on affective face processing. *Frontiers in Psychology*, *3*, 471. doi: org/10.3389/fpsyg.2012.00471

Wistedt, B., Rasmussen, A., Pedersen, L., Malm, U., Träskman-Bendz, L., Wakelin, J., & Bech, P. (1990). The development of an observer-scale for measuring social dysfunction and aggression. *Pharmacopsychiatry*, *23*, 249–252. doi: 10.1055/s-2007-1014514