Mindfulness, alexithymia, and empathy moderate relations between trait aggression and antisocial personality disorder traits

RUNNING HEAD: Mentalizing and Antisocial Personality Disorders

Patrizia Velotti [a], Carlo Garofalo [b], Giancarlo Dimaggio [c], Peter Fonagy [d]

[a] Educational Sciences Department, Genoa University, Corso Andrea Podestà, 2, 16128, Genoa, Italy. E-mail: patrizia.velotti@unige.it

ORCID ID: 0000-0002-1933-8314

[b] Department of Developmental Psychology, Tilburg University, Tilburg, the Netherlands. E-mail: c.garofalo@uvt.nl. ORCID: 0000-0003-2306-6961

[c] Center for Metacognitive Interpersonal Therapy, Piazza dei Martiri di Belfiore, 4, 00195, Rome, Italy. E-mail: gdimaje@libero.it

[d] Psychoanalysis Unit, Research Department of Clinical, Educational and Health Psychology University College London, Gower Street, London WC1E 6BT. E-mail: p.fonagy@ucl.ac.uk

Correspondence concerning this article should be addressed to Patrizia Velotti, Department of Educational Sciences, University of Genoa, Corso Andrea Podestà, 2, 16128, Genoa (Italy). Telephone: +39 010 20953722. E-mail: patrizia.velotti@unige.it

Conflict of Interest: The authors declare that they have no conflict of interest.

Mindfulness, alexithymia, and empathy moderate relations between trait aggression and antisocial personality disorder traits

Abstract

Antisocial personality disorder (ASPD) has long been described focusing exclusively on behavioral features, like aggression. Although the role of mentalizing for aggression is well established, research on the role of mentalizing in ASPD remains limited. The present study examined the independent and interactive effects of mentalizing abilities and aggression in predicting ASPD traits in a violent male offender sample (*N*=403). Participants completed self-report measures of ASPD traits, and a comprehensive assessment of mentalizing skills including measures of mindfulness, empathy, and alexithymia. Above and beyond the main effect of aggression, mindfulness, alexithymia, and empathy significantly explained an incremental amount of variance in ASPD traits in separate regression analyses. Further, mindfulness, alexithymia, and empathy significantly interacted with aggression in predicting ASPD scores. Findings suggest that among offenders with better mentalizing, aggression was significantly more strongly related to ASPD traits, indicating that at higher levels of mentalizing, only participants who also had higher levels of aggression scored higher on ASPD traits. Conversely, among offenders with poorer mentalizing, the positive association between aggression and ASPD traits was significantly weaker, indicating that poor mentalizing alone was sufficient to have high levels of ASPD traits (and aggression). Findings suggest that in some instances mentalizing may not have a protective role on ASPD in the presence of very high levels of aggression. However, among offenders with poor mentalizing, treatments targeting aggression may not be successful in reducing ASPD traits. Interventions that aim at improving mentalizing may ultimately be more effective to treat aggression and ASPD among offenders.

Keywords: Mentalizing, Mindfulness, Alexithymia, Aggression, Antisocial Personality Disorder, Offenders

Mindfulness, alexithymia, and empathy moderate relations between trait aggression and antisocial personality disorder traits

Individuals with antisocial personality disorder (ASPD) are prone to disrespect social norms; they may take advantage of others, manipulate, deceive, intimidate other people, and often manifest violent behavior (American Psychiatric Association 2013). They are at risk for many adverse outcomes, including anxiety disorders (Goodwin and Hamilton 2003), suicide (Zaheer et al. 2008), and addictive behaviors (Compton et al. 2005). The literature on ASPD highlights that ASPD patients have difficulties in controlling their impulses and often disregard the consequences of their actions (Swann et al. 2009). As result, they also pose a problem for the community, as they often enact domestic violence and other forms of aggression (Black 2015; Ullrich and Coid 2009; Velotti et al. 2018). ASPD occurs with a prevalence of 1–4% in the general population (Trull et al. 2010; Werner et al. 2015), with higher rates (50–60%) being found in inmate populations (Black et al. 2010; Ogloff 2006). Overall, ASPD places a substantial burden on society, both in terms of direct costs (e.g., incarceration, hospitalization) and through costs associated with the consequences of antisocial behavior for its victims (Bateman et al. 2013). Moreover, once these persons are arrested, they are difficult to manage because of their irritability and aggressiveness (Black et al. 2010). Thus, understanding the psychological mechanisms underlying ASPD is necessary in order to ameliorate their adjustment to prison life and to tailor targeted treatments that could contribute to a reduction of ASPD symptomology and its consequences.

In recent years, it has been proposed that an important factor to understand ASPD is poor mentalizing, which in turn may represent a candidate treatment target in this population (Bateman et al. 2013). Mentalizing refers to the mental processes by which humans reflect upon mental states in order to make sense of the actions and behavior of both themselves and

others (Fonagy 1991). Bateman and Fonagy (2004) described mentalizing as a "process by which an individual implicitly and explicitly interprets the actions of herself and others as meaningful on the basis of intentional mental states such as personal desires, needs, feelings, beliefs, and reasons" (p. xxi).

In keeping with Fonagy and Luyten (2009) and Lieberman (2007), mentalizing entails both affective and cognitive elements, and it can be focused on one's own or others' mental states. Indeed, many elements constitute the mentalizing network. Some of them are highly specific, for example being able to understand the others' emotions from the eyes (Baron-Cohen et al. 2015). Other elements involve information from multiple channels, such as integrating multiple ideas of self and others in a coherent picture (Semerari et al. 2007). Generally, mentalizing refers to a range of different operations, including: being aware of what one feels; being mindful, that is, being aware of what one experiences in a nonjudgmental way and without acting out; and understanding the mental states of others while resonating with them, that is, being empathic.

From this perspective, mentalizing is clearly a multifaceted ability (Fonagy and Luyten 2009; Semerari et al. 2007) that may be better assessed focusing on its subcomponents (Choi-Kain and Gunderson 2008). Mindfulness, alexithymia, and empathy appear to be each an part of the larger mentalizing network. Mindfulness is the capacity to live the present experience adopting a non-judgmental, accepting stance (Kabat-Zinn 2003) and is a form of cognitive reflection, inwardly directed and focused on both cognitions and affects. Alexithymia refers to difficulties identifying, understanding, and describing emotions, difficulties distinguishing feelings from sensations of emotional arousal, limited imaginal processes, and an externally oriented style of thinking (Bagby and Taylor 1997). Empathy is "the process by which an individual infers the affective state of another by generating an isomorphic affective state in the self, while retaining knowledge that the cause of the affective

state is the other" (Engen and Singer 2013, p. 277). All these processes have in common a focus on mental states, which makes them belong to the mentalizing network of abilities. However, they are also distinct enough to tap on different aspects of mentalizing. In turn, they each may have different links with aggression and antisocial tendencies.

As regard mindfulness, some individuals may fail to reflect on their inner reactions without acting, so to abstain from maladaptive and impulsive actions which they use to soothe distressing inner states. Indeed, there is evidence that poor mindfulness is related to greater aggressive tendencies in offenders (Velotti et al. 2016). Further, poor awareness of own affect, i.e. alexithymia, is a path to poor emotion regulation, which is a risk factor for aggression (Garofalo et al. 2018; Taylor et al. 1997). With poor awareness of what one feels, the individual under distress may lack information about what kind of affect he or she is experiencing and how to best soothe it (Velotti et al 2017). Finally, individuals with low empathy do not have interest or capacity to realize that others may suffer from the consequences of their action, and therefore abstain from aggression (Lovett and Sheffield 2007).

Mentalizing (and the related construct of metacognition; Carcione et al. 2011; Dimaggio and Lysaker 2010; Semerari et al. 2003) has been extensively studied in relation to personality disorders (e.g., Bateman and Fonagy 2004; Bo et al. 2014). However, only few studies have tested the assumption that poor mentalizing characterizes ASPD (Bateman et al. 2013; Beeney et al. 2015; McGauley et al. 2011), and virtually no studies have examined the relative contribution of and possible interaction between mentalizing skills and aggression in relation to ASPD traits. Indirect evidence for poor mentalizing skills in individuals with ASPD traits was provided by a recent study that showed that ASPD traits were significantly related to hostile interpretation bias in response to scenarios describing unintentionally provocative situations (Lobbestael et al. 2013). In addition, offenders with ASPD have been

found to perform poorly when asked to talk about their emotional experiences, such that their narratives were inaccurate in describing the emotions they felt, thus pointing to the possibility that ASPD is related to alexithymic features (Gawda 2011).

In a seminal study, Dolan and Fullam (2004) compared a sample of individuals meeting diagnostic criteria for ASPD and psychopathy, individuals meeting criteria for ASPD only, and a control group of non-clinical participants on basic and complex theory of mind tests. Findings from this study revealed that ASPD patients did not show deficits in basic theory of mind, but performed worse than controls on subtle tests, such as the faux pas test. Of note, these results were not qualified by the presence of psychopathy. Furthermore, ASPD-only patients showed impairments in facial emotion recognition, compared with both ASPD patients with comorbid psychopathy and controls (Dolan and Fullam 2004). Recent studies have highlighted volumetric reductions in the brains of individuals with ASPD, affecting areas involved in the processing of self-referential information and recognizing emotions of others (Bertsch et al. 2013). In addition, disrupted affective processing has been linked to the lack of empathy that people with ASPD often show (Zafirakis 2009).

Mindfulness deficits could also be related to ASPD. With low mindfulness, a person may not be able to take distance from thoughts that elicited anger or the will to attack others, and so there are no buffers preventing antisocial tendencies. As regards evidence for a correlation between reduced mindfulness and antisocial features, Fossati et al. (2012) found that mindfulness deficits were related to heightened ASPD traits in a clinical sample. This finding was recently replicated in a sample of violent offenders, in which significant negative correlations were found between mindfulness abilities (and specifically the ability to act with awareness) and self-reported ASPD traits (Velotti et al. 2016).

Finally, a recent study on mentalizing skills included items items tapping into emotional understanding, empathy, and mindfulness. Among community-dwelling

individuals, a positive association was found between scores on this index of poor mentalizing and clinician-rated ASPD traits, even after controlling for borderline personality disorder symptoms (Beeney et al. 2015). However, only 9 out of 152 participants met diagnostic criteria for ASPD, calling for further replication in samples with a greater prevalence of ASPD traits.

Summarizing, all these three aspects of the mentalizing network at stake here, may be the among the roots of antisocial behavior (see also Misso et al. 2018). Individuals that are less capable to mindfully regulate distressing states, that are less able to describe what they feel, and have diminished empathy might be prone to act in ways that harm others and disregard social norms and expectations. However, it is unclear whether these mentalizing deficits (i.e. low mindfulness, alexithymia, and diminished empathy) are defining features of ASPD, or whether their relevance for ASPD is limited to their role as risk factors for the behavioral aggressive tendencies typically related to ASPD. In addition, it is unclear whether both poor mentalizing and aggression independently contribute to ASPD traits.

Recently, Velotti et al. (2016) examined this possibility and found that mindfulness interacted with aggression in predicting ASPD. Low mindfulness was linked to more ASPD traits in a sample of violent offenders. Moreover, in offenders with low mindfulness, aggression did not predict antisocial traits. This likely means that mentalizing deficits, at least as regard mindfulness, may be sufficient to predict greater levels of ASPD, whereas the role of trait aggression is less relevant in this context. This is not to say that those participants showed low levels of aggression; actually, offenders with low mindfulness skills showed high levels of aggression and ASPD traits, but increases in aggression were not associated with increases in ASPD traits (Velotti et al. 2016). However, Velotti et al.'s (2016) study was limited by a relatively small sample size (N = 83) and by the inclusion of only one domain of the broader mentalizing construct, that is, mindfulness (Fonagy and Bateman 2016).

With these considerations in mind, the aim of the present was first to replicate findings from Velotti et al. (2016) in a larger sample of violent offenders, testing the moderating role of mindfulness in the relation between trait aggression and ASPD traits. Further, we explored the interaction between both alexithymia and empathy and aggression in predicting ASPD traits. We hypothesized that all these aspects of mentalizingm and aggression predicted ASPD traits. Further, we expected these mentalizing skills to interact with aggression in the same way as found by Velotti et al. (2016) concerning mindfulness: that is, among individuals with poor mentalizing, we expected the association between aggressive tendencies and ASPD to be non-significant.

Method

Participants

The sample comprised 403 male inmates recruited from different prisons in Northern Italy. All inmates were serving sentences for violent offending (i.e., they had committed crimes involving physical violence toward others, such as armed robbery, assault, homicide, sexual abuse, etc.). Mean age was 39.91 years (range: 19–77; *SD* = 11.79). Exclusion criteria were the presence of any psychotic disorder at the time of the study and drug or alcohol intoxication in the previous 3 months.

Procedure

All participants received a complete description of the study and signed written informed consent before completing the measures described below. Participation was voluntary and confidential, and participants did not receive any compensation. Inmates were also assured that their decision to take part in the study would not have any impact on their inmate status.

The assessment took place in individual or small-group sessions in the presence of a trained clinical psychologist, who helped participants in understanding items only upon request.

Measures

Antisocial Personality Disorder Traits. ASPD traits were assessed using the ASPD scale of the Italian version of the Millon Clinical Multiaxial Inventory–III (MCMI–III) (Millon 2006). The MCMI–III is a 175-item self-report instrument, which assesses dimensional scores of 14 personality styles and 10 clinical syndromes according to Millon's personality theory (2006) and to criteria from the Diagnostic and Statistical Manual of Mental Disorders (4th ed., text rev.; DSM-IV-T; American Psychiatric Association, 2000). ASPD is defined by traditional indicators such as a history of truancy and delinquency, as well as by more general antisocial traits and tendencies (e.g., being intimidating, dominating, competitive, self-reliant, vengeful, fearless, angry or hostile). Specifically, 7 items directly addressing prototypical features of ASPD are weighted 2, whereas 10 items, relating to less defining—although often associated—characteristics of ASPD (e.g., substance abuse, or a derogation of intimacy, warmth, and gentleness as sign of weakness) are weighted 1.

Trait Aggression. The Aggression Questionnaire (AQ; Buss and Perry 1992) was used to assess trait aggression. The AQ is a 29-item multidimensional self-report scale, which measures aggression through four subscales: physical aggression; verbal aggression; anger; and hostility. According to Buss and Perry's (1992) conceptualization of trait aggression, physical and verbal aggression represent the instrumental component, anger measures the affective component, and hostility captures the cognitive component of the overall aggression construct, which is indexed by the AQ total score. Respondents are asked to indicate how much each item applies to them on a 5-point Likert scale, with higher scores corresponding to

greater trait aggression. Both the original and the Italian version (Fossati et al. 2003) of the AQ demonstrated good psychometric properties.

Alexithymia. Emotional understanding was also assessed with the 20-item Toronto Alexithymia Scale-Revised (TAS-20; Bagby et al. 1994) total score. The TAS-20 is a self-report instrument that comprises 20 items rated on a 5-point Likert scale. The TAS-20 total score (with higher scores indicating greater alexithymia) is computed by summing scores on three dimensions: difficulty in identifying feelings; difficulty in describing feelings; and external-oriented thinking style. Since the psychometric properties of the subscales are questionable (Kooiman et al. 2002), we used the total score only as an index of overall alexithymia. The Italian version of the TAS-20 demonstrated good estimates of internal reliability (α =0.75; Bressi et al. 1996).

Mindfulness. Mindfulness levels were assessed using the Five Facet Mindfulness Questionnaire (FFMQ; Baer et al. 2006). The FFMQ is a 39-item self-report questionnaire rated on a 5-point Likert scale, measuring the respondent's ability to: attend to internal/external stimuli and associated cognitions and emotions; label and describe inner experiences with words; pay ongoing attention to present activities, with an associated awareness of personal motives behind one's behavior; adopt a non-evaluative stance towards one's thoughts and feelings; and perceive emotions and thoughts—especially if distressing—without feeling overwhelmed or compelled to react to them. The five subscales can be summed to produce a total score, with higher scores indicating greater mindfulness abilities. Validation of the Italian version of the FFMQ confirmed its good reliability and validity (Giovannini et al. 2014).

Empathy. Levels of empathy were assessed using the Empathy Quotient (EQ; Baron-Cohen and Wheelwright 2004). The questionnaire consists of 40 statements rated on a 4-point Likert scale. Scores can range from 0 to 80. The EQ seems to show acceptable internal consistency, concurrent and convergent validity, and good test–retest reliability (Allison et al. 2011; Preti et al. 2011) and in this study Cronbach's alpha was 0.82.

Data Analyses

Descriptive statistics were examined and Pearson's product-moment correlations were computed to test bivariate relations among all study variables. Next, hierarchical multiple regression analyses were carried out to investigate the independent and unique contribution of aggression and mentalizing dimensions (separately) to ASPD traits, in Step 1, as well as their interaction. Specifically, in Step 1 of each regression model, the ASPD dimensional score was regressed on the AQ total score and FFMQ total score, TAS 20-R total score, and EQ total score, respectively. In Step 2 of each regression model, the product-term of the AQ total score with each of the three mentalizing-related variables was introduced to test their interaction. Following Aiken and West's (1991) recommendations, predictor variables (i.e., AQ, FFMQ, TAS20, and EQ scales) were mean centered to compute the cross-product vector, which represented the interaction effect. Simple slope analyses were carried out to probe significant interaction effects at high and low levels of the moderators (i.e., 1 standard deviation above or below the mean, respectively).

Results

Descriptive statistics and internal consistency coefficients for all study variables are reported in Table 1. Table 1 also shows bivariate associations among all study variables. As would be expected on conceptual grounds, relatively stronger associations were reported among the

different variables used to capture mentalizing abilities or lack thereof (i.e., mindfulness, alexithymia, and empathy). In general, all variables were significantly intercorrelated, with effect sizes ranging between |0.26| (for the negative associations between empathy and ASPD traits) and |0.62| (for the negative relation between alexithymia and mindfulness).

[Insert Table 1 about here]

Table 2 shows results of the moderated regression analyses and simple slope analyses conducted to test and probe significant interaction effects between mentalizing components and aggression in predicting ASPD traits. Aggression was significantly and independently related to ASPD in all regression models, including the different components of mentalizing. The three regression models also revealed that, when controlling for aggression, there were significant main effects of mindfulness, alexithymia, and empathy on ASPD traits. Further, all three interaction effects were statistically significant, and the interaction terms were able to explain between 2 and 4% of additional variance in ASPD, above and beyond the main effect of aggression and each mentalizing component. Overall, the models explained between 23% (for the model including alexithymia and aggression) and 39% (for the model including mindfulness and aggression) of the variance in ASPD traits. Inspection of simple slope analyses revealed that the association between aggression and ASPD traits was significant and positive both at high and at low levels of the moderators. However, at low levels of mindfulness, both low and high aggression predicted ASPD, while at higher levels of mindfulness ASPD occurred only in the presence of high levels of aggression. Specifically, the relation between aggression and ASPD traits was stronger at higher levels of mindfulness (compared with lower levels of mindfulness) as well as at higher levels of empathy (compared with lower levels of empathy). Conversely, the relation between aggression and ASPD traits was stronger at lower levels of alexithymia (compared with higher levels of alexithymia). Taken together, these findings suggest that among individuals with better mentalizing skills

across the different domains (i.e., high mindfulness, high empathy and low alexithymia), aggression was significantly more strongly related to ASPD traits. In contrast, among individuals with poor mentalizing skills across the different domains (i.e., low mindfulness, low empathy and high alexithymia), aggression played a less relevant role in contributing to increased levels of ASPD traits. This is not to say that individuals with better mentalizing skills were also more likely to have higher levels of ASPD traits, but rather, mentalizing could not protect an individual from ASPD in the presence of very high levels of aggression. As can be seen in Figure 1 (which shows a graphical depiction of such an interaction effect, using the Aggression × Mindfulness interaction as an example), at lower levels of mentalizing skills (here, mindfulness), the level of ASPD traits was relatively higher than in individuals with higher levels of mentalizing/mindfulness, and comparatively less variability was explained by individual differences in aggression. Therefore, poor mentalizing seemed to be sufficient for individuals to have higher levels of ASPD traits. However, at higher levels of mentalizing/mindfulness, only participants who also had higher levels of aggression scored higher on ASPD traits, indicating that among those with better mentalizing skills, greater levels of aggressive tendencies were needed to yield higher ASPD trait scores[1].

[Insert Table 2 and Figure 1 around here]

Discussion

Among the possible reasons for antisocial behaviors there are trait aggressive tendencies , and poor capacity to understand the mental states of both oneself and others, as well as to regulate one's behavior on the basis of such awareness of own and others' thoughts and feelings. This capacity to recognize mental states and reason about them in order to pragmatically use them

[1] Follow-up exploratory analyses were conducted entering all predictors (and interaction terms) in the same model. This model showed that only aggression remained as a significant predictor of ASPD traits, suggesting that it may be the shared variance between empathy, mindfulness, and alexithymia that drives associations with ASPD traits and interacts with aggression in predicting ASPD traits.

during social interactions has been termed mentalizing (Bateman and Fonagy 2004) or metacognition (Semerari et al. 2003). Nevertheless, theories and descriptions of ASPD have typically emphasized the role of aggressive tendencies, neglecting the possible contribution of mentalizing skills. We hypothesized that different aspects of this capacity could interact with aggression in predicting ASPD traits. Our initial predictions, based on earlier findings by Velotti et al. (2016) were that mentalizing abilities and aggression dimensions were related to one another, and were both associated with ASPD traits. In addition, we expected that mentalizing skills and aggression interacted in predicting ASPD traits. Results largely supported these hypotheses. Confirming previous knowledge, we found a relation between aggression and ASPD. Furthermore, our findings corroborated the role of all three facets of mentalizing that we considered, that is mindfulness, alexithymia, and empathy, in predicting ASPD scores, above and beyond aggression. Moreover, these three variables exhibited a consistent pattern of interaction with aggression in predicting ASPD traits. Indeed, the relation between aggression and ASPD traits was stronger in individuals with high levels of mindfulness and empathy and low levels of alexithymia. In contrast, in individuals with poor mentalizing skills across the same three domains, aggression was more weakly related to ASPD traits.

Our findings suggest that these poor mentalizing skills are not only strongly related to ASPD traits, but in the presence of poor mentalizing skills, the role of trait aggression in predicting levels of ASPD becomes trivial. These findings replicate and extend Velotti et al.'s (2016) study. In this prior study, at low levels of mindfulness, the aggression–ASPD link became nonsignificant. The present study, involving a larger sample, indicates that this earlier finding may be due to limited statistical power of the study of Velotti et al. (2016), because the present findings show that aggression remained significantly and positively related to ASPD traits at low levels of mentalizing skills. Nevertheless, these relationship were

significantly weaker than those found among individuals with high levels of mentalizing. Therefore, in keeping with Velotti et al.'s (2016) conclusions, the findings of the present study suggest that when mentalizing skills are lacking, offenders score high on ASPD traits and aggression, and individual differences in aggression have minimal impact on ASPD traits. Extending Velotti et al.'s (2016) findings, we showed that this pattern does not apply only to mindfulness, but also involves other mentalizing domains such as alexithymia and empathy.

The most likely explanation is that for some individuals, the lack of capacity to understand the mental states of themselves and others, and to be mindful about them, paves the way to antisocial tendencies. If individuals are poorly aware of their own emotions, have a limited ability to grasp others' experience and so feel less concerned about them, and have poor capacity for mindfulness, it is likely that when they experience distress during social interactions they tend to react without thinking or calming themselves down. More generally, lacking awareness of their own mental states and those of others, these individuals may develop hostile and antagonistic tendencies and may tend to externalize blame for their misconduct. For example, when they feel threatened, humiliated, and abandoned, they may not be able to mindfully soothe themselves and be in touch with their own emotions; instead, they seek different ways to meet their needs, such as performing antisocial acts in both reactive and instrumental ways. In addition, those with poor empathy cannot take the perspective of the other and think that the other did not want to hurt, humiliate, or abandon them, nor they can be concerned about the suffering they are causing the other with their antisocial actions. Our findings revealed that individuals with poor mentalizing skills tended to also show aggressive tendencies, and variations in these aggressive tendencies had limited influence on the severity of ASPD traits. Therefore, treatments exclusively focused on reducing aggression would likely be unable to reduce ASPD symptoms, as long as

mentalizing skills are not improved. In contrast, interventions aimed at improving mentalizing may reduce both aggression and antisocial tendencies (Misso et al. 2018).

On the other hand, other individuals tend to enact antisocial behaviors in a predatory and premeditated way (Bateman and Fonagy 2011), regardless of their understanding of what they feel and what others think and feel, and irrespective of their capacity to live in the present moment, be aware of the experience, and mindfully attend to it. These individuals' behaviors are likely not led by their current inner state—that is, they do not act under the pressure of painful emotions and attack the other in order to soothe distress stemming from interpersonal sources. Some antisocial offenders can have intact mentalizing skills, and their levels of ASPD traits may be mostly dependent on their aggressive tendencies. For this group of offenders, treatments aimed at increasing mentalizing may not be needed or appropriate. Rather, interventions should be aimed at reducing the hostile attitudes that lead them to use mentalizing skills in antisocial ways rather than in the service of prosocial goals.

Limitations and Future Research Directions

There were limitations to our study. The choice to consider mindfulness, alexithymia, and empathy as part of a broader mentalization factor was exclusively based on conceptual grounds, and we did not examine their association with established measure of reflective functioning. The self-report nature of the assessment of mentalizing and the other variables bears the risk of inflated correlations due to shared method variance. Again, individuals may not have been correct in their descriptions of their own mentalizing difficulties, and so replication is urgently needed with laboratory measures and interviews to capture participants' ability to understand and regulate mental states. The cross-sectional design of our study is another limitation as it does not allow us to understand causal relations among

variables. Future studies adopting longitudinal or experimental designs are required to disentangle the temporal and causal relationships among the components of the mentalizing system, as well as their role in ASPD. Although exploring an offender population, where antisocial behavior is highly relevant, is important, this is also a limitation, as results cannot be generalized to other clinical or community samples displaying problems with aggression or subthreshold levels of ASPD traits. Replication is therefore needed with participants drawn from the community.

Despite these limitations, the present findings may have relevant implications for future research, being among the first empirical tests of the mentalizing framework for ASPD (Bateman et al. 2013). Assessing mentalizing may be key to planning new research designs to investigate ASPD in forensic and correctional settings. Considering that individuals with ASPD manifest behaviors marked by levels of aggression, both trait- and state-like, it could be useful for future research also to look at the underlying processes that may link ASPD traits and aggression considering aggression as main outcome.

Author Contributions

PV: designed and executed the study and wrote the paper. CG: analyzed the data, wrote the first draft of the results, and collaborated in the writing and editing of the final version of the manuscript. GD and PF: collaborated in the writing and editing the final version of manuscript.

Compliance with Ethical Standards

Conflict of Interest. The author(s) declares that they has no conflict of interest.

Ethical Approval. All procedures performed in this study involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Also, the study received formal approval both from the ethics committee of Sapienza University of Rome and the Italian Ministry of Justice.

Informed Consent Informed consent was obtained from all individual participants included in the study

References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park: Sage.

Allison, C., Baron-Cohen, S., Wheelwright, S. J., Stone, M. H., & Muncer, S. J. (2011). Psychometric analysis of the Empathy Quotient (EQ). *Personality and Individual Differences, 51*, 829-835. doi:10.1016/j.paid.2011.07.005.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (5th ed.).* Washington, DC: Author.

Baer, R. A., Smith, G. T., Hopkins, J., Krietemeyer, J., & Toney, L. (2006). Using self report assessment methods to explore facets of mindfulness. *Assessment, 13(1)*, 27-45. doi: 10.1177/1073191105283504.

Bagby, R. M., Parker, J. D. A., & Taylor, G. J. (1994). The 20-item Toronto alexithymia scale. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research, 38(1),* 23-32.

Bagby, R. M., & Taylor (1997). *Affect dysregulation and alexithymia, Disorders of affect regulation: alexithymia in medical and psychiatric illness,* Cambridge: University Press.

Baron-Cohen, S., Bowen, D. C., Holt, R. J., Allison, C., Auyeung, B., Lombardo, M. V., … Lai, M.-C. (2015). The "Reading the Mind in the Eyes" Test: Complete Absence of Typical Sex Difference in ~400 Men and Women with Autism. *PLoS ONE, 10*(8), e0136521. http://doi.org/10.1371/journal.pone.0136521.

Baron-Cohen, S., & Wheelwright, S. (2004). The Empathy Quotient (EQ): An investigation of adults with Asperger syndrome and high-functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders, 34,* 163-175.

Bateman, A., Bolton, R., & Fonagy, P. (2013). Antisocial personality disorder: A mentalizing framework. *Focus, 11(2),* 178-186.

Bateman, A., & Fonagy, P. (2004). *Psychotherapy for borderline personality disorder: Mentalization-based treatment*. Oxford: Oxford University Press.

Bateman, A., & Fonagy, P. (2011). Antisocial personality disorder. In A. Bateman & P. Fonagy (Eds.), *Handbook of mentalizing in mental health practice* (pp. 357-378). Washington DC: American Psychiatric Publishing, Inc.

Beeney J. E., Stepp S. D., Hallquist M. N., Scott L. N., Wright A. G. C., & Ellison W. D. (2015). Attachment and social cognition in borderline personality disorder: Specificity in relation to antisocial and avoidant personality disorders. *Personality Disorders: Theory, Research, and Treatment, 6,* 207-215. doi: 10.1037/per0000110.

Bertsch, K., Grothe, M., Prehn, K., Vohs, K., Berger, C., Hauenstein, K., et al. (2013). Brain volumes differ between diagnostic groups of violent criminal offenders. *European*

Archives of Psychiatry and Clinical Neuroscience, 263(7), 593-606. doi: 10.1007/s00406-013-0391-6.

Black, D. W. (2015). The natural history of antisocial personality disorder. *Canadian Journal of Psychiatry, 60(7),* 309-314. doi: 10.1177/070674371506000703.

Black, D. W., Gunter, T., Loveless, P., Allen, J., & Sieleni, B. (2010). Antisocial personality disorder in incarcerated offenders: Psychiatric comorbidity and quality of life. *Annals of Clinical Psychiatry, 22*, 113-120.

Bo, S., Abu-Akel, A., & Kongerslev, M.T. (2014). Metacognition as a framework to understanding the occurrence of aggression and violence in patients with schizophrenia. In P. H. Lysaker, M. Brune, & G. Dimaggio (Eds.), *Social cognition and metacognition in schizophrenia: Psychopathology and treatment approaches* (pp. 137-145). London: Academic Press.

Bressi, C., Taylor, G., Parker, J., Bressi, S., Brambilla, V., Aguglia, E., et al. (1996). Cross validation of the factor structure of the 20-item Toronto Alexithymia Scale: an Italian multicenter study. *Journal of Psychosomatic Research, 41(6),* 551-559.

Buss, A. H., & Perry, M. (1992). The Aggression Questionnaire. *Journal of Personality and Social Psychology, 63(3)*, 452-459.

Carcione, A., Nicolò, G., Pedone, R., Popolo, R., Conti, L., Fiore, D., … Dimaggio, G. (2011). Metacognitive mastery dysfunctions in personality disorder psychotherapy. *Psychiatry Research, 190,* 60-71. doi: 10.1016/j.psychres.2010.12.032.

Choi-Kain, L. W., & Gunderson, J. G. (2008). Mentalization: Ontogeny, assessment, and application in the treatment of borderline personality disorder. *American Journal of Psychiatry, 165(9),* 1127-1135. doi: 10.1176/appi.ajp.2008.07081360.

Compton, W. M., Conway, K. P., Stinson, F. S., Grant, B.F. (2005). Prevalence, correlates, and comorbidity of DSM-IV antisocial personality syndromes and specific drug use

disorders in the United States: Results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Journal of Clinical Psychiatry, 66,* 677-685.

Dimaggio, G., & Lysaker, P. H. (Eds.). (2010). *Metacognition and severe adult mental disorders: From basic research to treatment*. London: Routledge.

Dolan, M., & Fullam, R. (2004). Theory of mind and mentalizing ability in antisocial personality disorders with and without psychopathy. *Psychological Medicine, 34(6),* 1093-1102.

Engen, H. G., & Singer, T. (2013). Empathy circuits. *Current Opinion in Neurobiology, 23*(2), 275-282. doi: 10.1016/j.conb.2012.11.003.

Fonagy, P. (1991). Thinking about thinking: Some clinical and theoretical considerations in the treatment of a borderline patient. *International Journal of Psychoanalysis, 72,* 639–656.

Fonagy, P. (2003). The development of psychopathology from infancy to adulthood: The mysterious unfolding of disturbance in time. *Infant Mental Health Journal, 24(3),* 212-239. https://doi.org/10.1002/imhj.10053.

Fonagy, P., & Bateman, A. W. (2016). Adversity, attachment, and mentalizing. *Comprehensive Psychiatry, 64,* 59-66. doi: 10.1016/j.comppsych.2015.11.006.

Fonagy, P., & Luyten, P. (2009). A developmental, mentalization-based approach to the understanding and treatment of borderline personality disorder. *Development and Psychopathology, 21,* 1355-1381.

Fossati, A., Maffei, C., Acquarini, E., & Di Ceglie, A. (2003). Multigroup confirmatory component and factor analyses of the Italian version of the Aggression Questionnaire. *European Journal of Psychological Assessment, 19(1),* 54-65.

Fossati, A., Vigorelli Porro, F., Maffei, C., & Borroni, S. (2012). Are the DSM-IV personality disorders related to mindfulness? An Italian study on clinical participants. *Journal of Clinical Psychology, 68(6),* 672-683. doi: 10.1002/jclp.21848.

Garofalo, C., Velotti, P., & Zavattini (2017). Emotion regulation and aggression: The incremental contribution of alexithymia, impulsivity, and emotion dysregulation facets. *Psychology of Violence, 8*(4), 470–483. doi: 10.1037/vio0000141.

Gawda, B. (2013). The emotional lexicon of individuals diagnosed with antisocial personality disorder. *Journal of Psycholinguistic Research, 42(6),* 571-580. doi: 10.1007/s10936-012-9237-z.

Giovannini, C., Giromini, L., Bonalume, L., Tagini, A., Lang, M., & Amadei, G. (2014). The Italian five facet mindfulness questionnaire: A contribution to its validity and reliability. *Journal of Psychopathology and Behavioral Assessment, 36(3),* 415-423.

Goodwin, R.D., & Hamilton, S.P. (2003). Lifetime comorbidity of antisocial personality disorder and anxiety disorders among adults in the community. *Psychiatry Research, 117(2),* 159-166.

Kabat-Zinn, J. (2003). Mindfulness-based interventions in context: past, present, and future. *Clinical Psychology: Science and Practice, 10,* 144–156. doi:10.1093/clipsy/bpg016.

Kooiman, C. G., Spinhoven, P., & Trijsburg, R. W. (2002).The assessment of alexithymia: A critical review of the literature and a psychometric study of the TAS-20. *Journal of Psychosomatic Research, 53*, 1083-1090. doi: 10.1007/s10936-012-9237-z.

Lieberman, M. D. (2007). Social cognitive neuroscience: A review of core processes. *Annual Review of Psychology, 58*, 259-289. doi:10.1146/annurev.psych.58.110405.085654.

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(1),* 53-66. doi: 10.1521/pedi.2013.27.1.53.

Lovett, B.J., & Sheffield, R.A. (2007). Affective empathy deficits in aggressive children and adolescents: a critical review. *Clin Psychol Rev., 27*(1), 1-13.

.Marsh, A. A., & Blair, R. J. (2008). Deficits in facial affect recognition among antisocial populations: A meta-analysis. *Neuroscience and Biobehavioral Reviews, 32(3),* 454-465.

McGauley, G., Yakeley, J., Williams, A., & Bateman, A. (2011). Attachment, mentalization and antisocial personality disorder: The possible contribution of mentalization-based treatment. *European Journal of Psychotherapy, Counselling and Health, 13(4),* 371-393.

Millon, T. (2006). *MCMI-III Manual. 3rd edition*. Minneapolis: NCS Pearson. [Italian adaptation: Zennaro, A., Ferracuti, S., Lang, M., & Sanavio, E., Eds. Firenze: Giunti OS; 2008].

Misso, D., Schweitzer, R.D. & Dimaggio, G. (2018). Metacognition: A Potential Mechanism of Change in the Psychotherapy of Perpetrators of Domestic Violence. *Journal of Psychotherapy Integration.* Advance online publication. http://psycnet.apa.org/doi/10.1037/int0000111.

Ogloff, J. R. (2006). Psychopathy/antisocial personality disorder conundrum. *.Australian and New Zealand Journal of Psychiatry, 40(6-7),* 519-528. doi: 10.1080/j.1440-1614.2006.01834.x.

Preti, A., Vellante, M., Baron-Cohen, S., Zucca, G., Petretto, D. R., & Masala, C. (2011). The Empathy Quotient: £a cross-cultural comparison of the Italian version. *Cognitive Neuropsychiatry, 16(1),* 50-70. doi: 10.1080/13546801003790982.

Semerari A., Carcione A., Dimaggio G., Falcone M., Nicolo G., Procacci M., Alleva, G. (2003). How to evaluate metacognitive functioning in psychotherapy? The metacognition assessment scale and its applications. *Clinical Psychology & Psychotherapy, 10,* 238-261. https://doi.org/10.1002/cpp.362.

Semerari, A., Carcione, A., Dimaggio, G., Nicolò, G., & Procacci, M. (2007). Understanding minds, different functions and different disorders? The contribution of psychotherapeutic research. *Psychiatry Research, 17,* 106-119. doi: 10.1080/10503300500536953.

Swann, A. C., Lijffijt, M., Lane, S. D., Steinberg, J. L., & Moeller, F. G. (2009). Trait impulsivity and response inhibition in antisocial personality disorder. *Journal of Psychiatric Research, 43,* 1057-1063. doi: 10.1016/j.jpsychires.2009.03.003.

Taylor, G. J., Bagby, R. M., & Parker, J. D. A. (1997). *Disorder of affect regulation: Alexithymia in medical and psychiatric illness.* Cambridge: Cambridge University Press.

Taubner, S., White, L. O., Zimmermann, J., Fonagy, P., & Nolte, T. (2013). Attachment-related mentalization moderates the relationship between psychopathic traits and proactive aggression in adolescence. *Journal of Abnormal Child Psychology, 41(6),* 929-938.

Trull, T. J., Jahng, S., Tomko, R. L., Wood, P. K., & Sher, K. J. (2010). Revised NESARC personality disorder diagnoses: Gender, prevalence, and comorbidity with substance dependence disorders. *Journal of Personality Disorders*, *24(4*), 412-426. doi: 10.1521/pedi.2010.24.4.412.

Ullrich, & S., Coid, J. (2009). Antisocial personality disorder: Co-morbid Axis I mental disorders and health service use among a national household population. *Personality and Mental Health, 3,* 151-164. https://doi.org/10.1002/pmh.70.

Velotti, P., Beomonte Zobel, S., Rogier, G., Tambelli, R. (2018). Exploring relationships: A systematic Review on Intimate Partner Violence and Attachment, *Frontiers in Psychology, 9.* https://doi.org/10.3389/fpsyg.2018.01166.

Velotti, P., Garofalo, C. Callea, A., Bucks, R., Roberton, T. & Daffern, M. (2017). Exploring anger among offenders: The role of emotion dysregulation and alexithymia. *Psychiatry, Psychology and Law, 28(1),* 128-138. doi: 10.1080/13218719.2016.1164639.

Velotti, P., Garofalo, C., D'Aguanno, M., Popolo, R., Salvatore, G. & Dimaggio, G. (2016). Mindfulness moderates the relationship between aggression and antisocial personality disorder traits: Preliminary investigation with an offender sample. *Comprehensive Psychiatry, 64,* 38-45. doi: 10.1016/j.comppsych.2015.08.004.

Werner, K. B., Few, L. R., & Bucholz, K. K. (2015). Epidemiology, comorbidity, and behavioral genetics of antisocial personality disorder and psychopathy. *Psychiatric Annals, 45(4),* 195-199. doi: 10.3928/00485713-20150401-08.

Zafirakis, E. (2009). Examining empathy and perspective taking among high-risk and low-risk persistently antisocial and non antisocial young adults. *International Journal of Interdisciplinary Social Sciences, 4(7),* 193-204.

Zaheer, J., Links, P. S., & Liu, E. (2008). Assessment and emergency management of suicidality in personality disorders. *Psychiatric Clinics of North America, 31*(3), 527-543. doi: 10.1016/j.psc.2008.03.007.

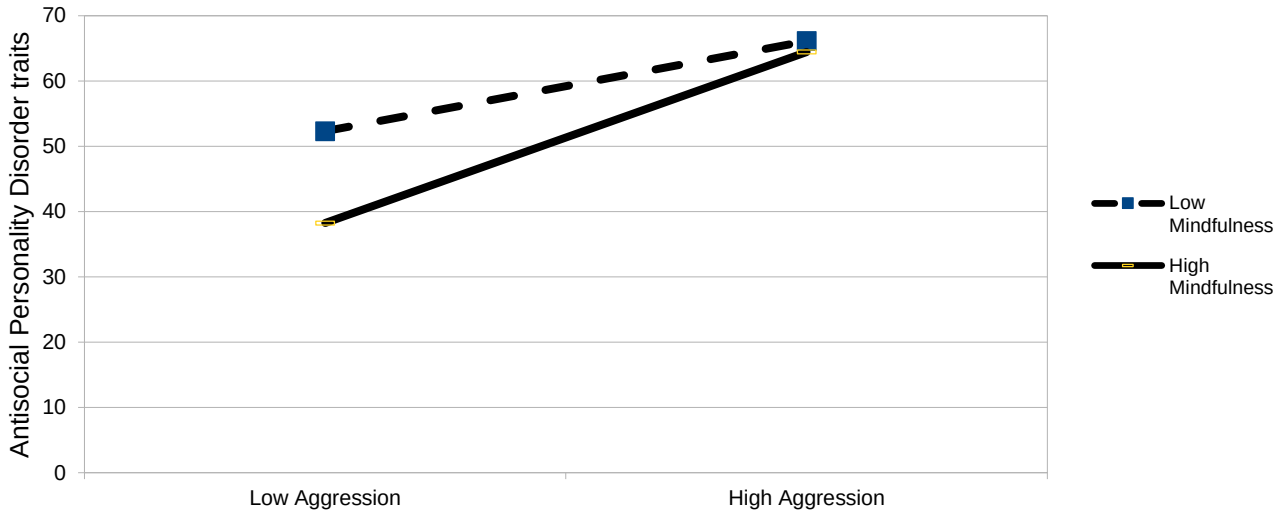

Figure 1. Graphical depiction of the significant interaction effect between mindfulness and trait aggression in predicting antisocial personality disorder traits. At low levels of mindfulness (dashed line), the slope is non-significant (i.e., aggression is not significantly related to antisocial personality disorder traits). At high levels of mindfulness (solid line), the slope is significant (i.e., increases in aggression scores are significantly related to increases in antisocial personality disorder traits).

Table 1

Internal consistencies, mean, standard deviations (SD), and bivariate associations for all study variables (total N = 403).

	α	*N (valid)*	*Range*	*M*	*SD*	1.	2.	3.	4.	5.
1. Antisocial Personality Disorder Traits	0.84	387	0 – 100	54.79	21.97	—				
2. Mindfulness	0.77	374	91 – 175	130.92	15.01	–0.32***	—			
3. Alexithymia	0.77	390	21 – 79	46.87	12.34	0.36***	–0.62***	—		
4. Empathy	0.83	379	15.14 – 70	40.52	10.67	–0.26***	0.47***	–0.45***	—	
5. Aggression	0.87	401	33 – 134	71.89	18.36	0.48***	–0.35***	0.43***	–0.27***	—

Note. * $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$.

Table 2

Multiple regression and simple slopes analyses testing the moderating effect of mindfulness, alexithymia, and empathy in the associations between aggression and antisocial personality disorder traits (N = 403; 5,000 bootstraps).

Dependent variable: ASPD traits					
	B (SE)	*95% CI*		*B (SE)*	*95% CI*
Intercept	55.27 (1.05)	53.20 to 57.34	Intercept	55.22 (1.04)	53.17 to 57.28
Aggression	0.56 (0.06)	0.44 to 0.63	Aggression	0.55 (0.06)	0.43 to 0.66
Mindfulness	–0.26 (0.07)	–0.40 to –0.12	Empathy	–0.25 (0.10)	–0.44 to –0.05
Aggression × Mindfulness	0.01 (0.003)	0.004 to 0.02	Aggression × Empathy	0.02 (0.01)	0.01 to 0.03
$R^2 = 0.29^{***}$; $\Delta R^2 = 0.02^{**}$			$R^2 = 0.27^{***}$; $\Delta R^2 = 0.02^{**}$		
Conditional effect of Aggression on ASPD traits at:					
High levels of Mindfulness	0.73 (0.09)	0.56 to 0.90	High levels of Empathy	0.72 (0.08)	0.57 to 0.89
Low levels of Mindfulness	0.38 (0.08)	0.26 to 0.53	Low levels of Empathy	0.37 (0.08)	0.21 to 0.52
Dependent variable: ASPD traits					
Intercept	56.38 (1.03)	54.35 to 58.42			
Aggression	0.55 (0.06)	0.43 to 0.66			
Alexithymia	0.35 (0.09)	0.18 to 0.52			
Aggression × Alexithymia	–0.01 (0.003)	–0.02 to –0.01			
$R^2 = 0.23^{***}$; $\Delta R^2 = 0.03^{***}$					
Conditional effect of Aggression on ASPD traits at:					
High levels of Alexithymia	0.36 (0.07)	0.23 to 0.50			
Low levels of Alexithymia	0.73 (0.08)	0.57 to 0.89			

Note. ASPD = Antisocial Personality Disorder. CI = Confidence Interval. High and low levels of the moderators refer to scores 1 standard deviation above or below the mean, respectively. $\Delta R^2 = R^2$ change due to the addition of the interaction term in the model (i.e., proportion of incremental variance explained by the interaction term above and beyond the main effects).

* $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$.