

**Confusing my viewpoint with his: Altered self-other distinction performance in antisocial
personality disorder**

June 1st, 2023

Abstract

Deficits of social cognition are regularly but inconsistently reported among individuals with antisocial personality disorder (ASPD). Because of the multifaceted nature of social cognition, deficits might be only observed when assessing specific facets of social cognition and under sufficiently demanding conditions. This study examined self-other distinction performance, a key facet lying at the core of the attachment-based model of mentalizing (Fonagy & Luyten, 2009). Twenty-one forensic inpatients with ASPD and 19 participants from the community completed a visual perspective-taking paradigm allowing to tease apart self-other priority (i.e., how self-focused one is) from self-other distinction performance (i.e., how much one confuses his with others' mental states). The ASPD group made significantly more errors at handling conflicting self-other viewpoints by enforcing self-other distinction (19%) than the control group (4%), but the ASPD group was not significantly more self-focused. In contrast, the Interpersonal Reactivity Index self-report scale did not differentiate the two groups. However, a novel measure of self-other distinction based on Empathic concern (i.e., the tendency to experience feelings of concern and compassion for others) and Personal distress subscales (i.e., the tendency to experience personal distress in response to the distress of others) did differentiate the two groups, albeit to a significantly lower extent than the objective measure of self-other distinction. Altogether, these findings indicate the presence of a self-other distinction deficit in ASPD and advocate for psychometric approaches that embrace the multifaceted nature of social cognition and the need for objective measures with sufficient sensitivity.

Keywords: Self-Other Distinction, Social Cognition, Antisocial personality Disorder, Perspective-Taking, Empathy, Theory of Mind, Mentalizing

Introduction

People with antisocial personality disorder (ASPD) suffer from poor social functioning that is mainly reflected by their antisocial behaviors such as high impulsivity and a low threshold for irritability and aggression, which place them at risk for various detrimental life outcomes (e.g., domestic violence, incarceration or hospitalization; Black, 2010; Trull et al., 2010; Werner et al., 2015). These maladjusted behaviors are accountable for significant collateral costs because of the legal procedures involved, but also because of damage and reparations to the victims and society (Bateman et al., 2013).

Aiming to understand and tackle the underlying causes of their poor social functioning, many studies have examined social cognition in ASPD, with a large majority pointing towards the existence of one or more deficient socio-cognitive processes, such as empathy, awareness of own emotions, emotions inhibition, or affective perspective-taking (Malterer et al., 2008; Van Dongen, 2020; Velotti et al., 2019; Yoder et al., 2022). However, despite the increasing body of research investigating distinct facets of social cognition as candidates for the ASPD socio-cognitive deficiencies, previous studies have failed to clearly define which specific facet of social cognition is impaired in individuals with ASPD. For instance, a recent systematic review conducted by Marsden et al. (2019) on studies investigating samples of individuals with ASPD, with or without comorbid psychopathy, found no conclusive evidence of empathy deficits in this population. These results contrast with two other meta-analyses (Campos et al., 2022; Miller & Eisenberg, 1988), that revealed the presence of both lower affective and cognitive empathy in ASPD, with a larger deficit for cognitive empathy. Both these reviews comprised studies using self-reported measures and objective scores. Indirect evidence from two meta-analyses (Jolliffe & Farrington, 2004; van Langen et al., 2014) reported strong associations between reduced

cognitive empathy and offending but these studies focused exclusively on samples officially characterized with delinquency (and not simply problem behaviors) who completed self-report assessment tools such as the Interpersonal Reactivity Index (Davis, 1983), the Hogan's Empathy Scale (Hogan, 1969) and the Questionnaire Measure of Emotional Empathy (Mehrabian & Epstein, 1972). However, eight out of the 38 studies analyzed by van Langen et al. (2014) failed to report an association between cognitive empathy and offending individuals. Others have used objective measures of social cognition, such as the Reading the Mind in the Eyes task (RMET; Baron-Cohen et al., 2001), which is frequently used to assess socio-cognitive abilities in a performance-based format. However, a recent meta-analysis conducted by Johnson et al. (2022) reported no significant association between the RMET and ASPD. Despite being frequently used to measure cognitive and affective mentalizing (or Theory of Mind), this task has been criticized regarding its suitability for such assessment due to its major emotional processing component (Eddy, 2019; Kittel et al., 2022; Oakley et al., 2016; Quesque & Rossetti, 2020). Finally, Chang et al., 2021), in their review, argued that the inconsistencies observed in ASPD research might be due to a lack of systematic identification and labeling of the precise components of cognitive empathy that are measured (e.g., implicit vs. explicit). Moreover, the authors state that the tasks used to study ASPDs' socio-cognitive deficits must be selected after careful consideration of the biases that could be elicited by the settings (for instance, in-group and out-group effects in the Movie for the Assessment of Social Cognition; MASC; Newbury-Helps et al., 2017). Further, while we acknowledge that the failure to establish a clear pattern of deficits may be partly caused by the high heterogeneity of profiles within the ASPD population (leading to different results from one sample to another), this study aims to examine the hypothesis that a double psychometric problem may be the cause as well. We posit that identifying precisely and

consistently the specific impairments in this population has remained highly challenging for two reasons.

The first reason relates to the failure to *specifically* measure the *right* facet of social cognition. A robust obstacle encountered in ASPD research has been the lack of construct clarity in social cognition, leading researchers to target too broadly defined facets and to incorrectly assume they measure similar/different facets of social cognition because they have the same/different labels (the latter problem is also known as the jingle-jangle fallacies; Lilienfeld & Strother, 2020). The consequences are particularly visible at the operationalization level: very few measurement tools have been designed to precisely unpack and assess the distinct facets that are at play (Sulzer et al., 2016). For instance, as highlighted by Hall and Schwartz (2019) in their review, a large majority of instruments mobilized when aiming to investigate empathy as a construct reflected more a combination of socio-cognitive facets (e.g., emotional contagion, perspective taking, empathic behavior) with little regard for their disentanglement. As a result, only limited or erroneous conclusions can be drawn as to the key facets at play. To illustrate, the IRI is the most widely used assessment tool in empathy and ASPD research and it is a self-report scale designed to measure the “reactions of one individual to the observed experiences of another”. IRI has four subscales: Fantasy, Perspective-taking (PT), Empathic concern (EC), and Personal distress (PD). A focus on the IRI-PT subscale, which is typically used as a proxy of one’s ability to adopt the point of view of the others, is frequently used in clinical populations. However, examination of the 7 items comprising the PT score reveals that this subscale primarily captures the *motivation to try* to take the others’ perspective (Murphy et al., 2022) in *prosocial* contexts rather than a tangible competence in a neutral context (e.g., “I sometimes try to understand my friends better by imagining how things look from their perspective”; Jolliffe &

1 Farrington, 2004; van Langen et al., 2014). In other words, the PT subscale is not informative
2 about how successful one is at taking the perspective of others but rather about the adherence to
3 socially-desirable communal values (x under review, 2024; see Table 1 of Supplemental Material
4 for a list of the items).

5 A second reason for the past inconsistencies in characterizing ASPD social cognition
6 relates to the adequacy of the format of the assessment tool. Indeed, the lack of construct clarity,
7 but also the pragmatic need to have practical measures of the socio-cognitive facets have led to
8 an operationalization in formats that are misaligned with the targeted construct (e.g., Oakley et
9 al., 2016). In line with Hall & Schwartz (2019)'s conclusions regarding socio-cognitive
10 assessment, we posit that one part of the problem stems from the predominant use of self-report
11 measures to assess ASPD's abilities. This format requires that participants be able to reflect on
12 their own condition or behaviors in a relatively detached and objective fashion (Neumann &
13 Westbury, 2011). Such requirements are prone to be highly influenced by various intrinsic
14 processes such as social desirability and metacognitive abilities (or insight, i.e., the capacity to
15 accurately evaluate one's own functioning; Dang et al., 2020) and by the accuracy of self-
16 knowledge, which is known to be biased in various ways (for a short review, see Bukowski,
17 2019). These limitations are aggravated in ASPD because this population is associated with poor
18 metacognitive abilities (Cankaya et al., 2022), high levels of narcissism (Paulhus & Williams,
19 2002), and reduced awareness of own emotions (Velotti et al., 2019). However, objective
20 behavioral measures may also present some limitations: on the one hand, they often measure
21 multiple and undifferentiated processes, and, on the other, they do not offer a sufficiently high
22 level of difficulty to objectify more subtle deficits. Specifically, we postulate that some tasks are
23 insufficiently demanding to detect socio-cognitive deficits in ASPD. In other words, we

hypothesize that a necessary degree of task stringency is needed to capture and highlight specific socio-cognitive deficits. Dolan and Fullam (2004), in their study, have shown that individuals with ASPD and without psychopathic traits did not significantly performed worse on basic tests of ToM, using Stone et al. (1998)’s first and second order trials and on the Faux Pas task. Only more subtle components of this task objectified a group difference (on the listener’s mental state and empathic understanding questions). Moreover, results showed that ASPD with psychopathy could read basic or complex emotions (= “mental states”) from facial expression as well, if not better, as healthy controls in the Facial emotional expression task (Baron-Cohen et al., 2001).

To recap, our two hypothesized explanations of the past inconsistencies in characterizing ASPD social-cognition deficits are the failure to target specifically the right facet of social cognition – H1 – and the failure to use the appropriate format to capture socio-cognitive deficits in ASPD – H2. In this study, we aim to address and test the two hypothesized issues by predicting first that measuring specifically *self-other distinction* (SOD), a narrower and lower-level facet of social cognition, will better differentiate ASPD individuals from healthy adults than the use of tasks that only partly measure SOD. Our second prediction is that measuring SOD with a sufficiently demanding behavioral task, which captures an objective performance in a neutral context, will better differentiate ASPD individuals from healthy adults than with a self-report questionnaire (see Table 2 of Supplemental Material for a recap).

Prediction 1: Assessing SOD specifically better differentiates ASPDs from controls

Luyten et al. (2021) present a compelling approach to socio-cognitive functioning in psychopathologies based on Bateman et al. (2013)’s mentalizing framework. The model posits that successfully engaging in mentalizing is underpinned by four dimensions: automatic/controlled, internally/ externally based, self/other oriented, and cognitive/affective

processing. According to the authors, imbalances stemming from the different dimensional levels may be of relevance to characterize various types of personality disorders, and particularly the self/other dimension. Indeed, imbalances within this dimension are manifested by symptoms that are thought to be consequences of impaired SOD. Failure to distinguish self-generated sensations, actions, emotions, and thoughts from those elicited by the external environment and other people leads to self-other confusions such as identity diffusion, feelings of not being in control or in possession of our body, attributing self-thoughts to others (e.g. hostility attribution), perceiving sensations as externally generated (e.g., hearing voices). Further, failure to flexibly focus on (or select) the self or other mental state that is most relevant in a given context leads individuals to rigid egocentricity (self-focus), making them blind to social signals and alternative viewpoints. In other contexts, however, these individuals who fail to concurrently maintain separate self and others' representations may be excessively influenced by their social environment and fear to lose themselves in social relationships (Luyten et al., 2021). Hence, this demonstrates that impaired SOD has diverse detrimental consequences on social functioning. Among these, Eddy (2018) also suggested that the externalized behaviors observed notably in ASPD individuals may in fact be a coping mechanism for the self and other confusions they experience (cf. Discussion). Although we acknowledge that SOD is unlikely to be the sole impairment in ASPD, results from previous studies seem to support this idea. Indeed, in line with the self-report biases previously highlighted, several self-related processes tend to be deficient in this population and could reflect disturbances on the self-other dimension. Beyond poor metacognitive abilities, narcissism, and lack of emotional awareness, ASPD individuals show reduced performance when attempting to generate self-defining memories. On the one hand, Vanderveren et al. (2021) found that antisocial symptoms were associated with narrative

1 incoherence when evoking a meaningful autobiographical memory. On the other hand, Lavallee
2 et al. (2020) found a lower integration of past experiences when asked to retrieve specific events
3 from their life history. According to the authors, these results are indicative of identity
4 disturbance and poor self-coherence in this population, thus supporting the assumption of a
5 disruption in SOD.

6 To our knowledge, no study has ever examined SOD abilities in ASPD with a measure
7 specifically targeting this facet of social cognition despite its central theoretical importance in
8 personality disorders. Nevertheless, a few measures do partly measure SOD, including the IRI.

9 **Assessment of SOD via the Interpersonal Reactivity Index**

10 The Interpersonal Reactivity Index is divided into 4 subscales sharing a latent construct
11 which is the “responsivity to others”. One of the subscales is labeled Personal distress (PD), and
12 is often defined as “the individual's own feelings of fear, apprehension, and discomfort at
13 witnessing the negative experiences of others”. PD is mechanistically understood as an
14 uncontrolled emotional contagion, as it encompasses a mixture of high sensitivity to the
15 emotional states of others, and/or a failure to dissociate one’s own emotional state from the state
16 of the other person. PD is therefore considered as a manifestation of poor or absent SOD, in the
17 sense that the emotional contagion has not been prevented, leading the individual to experience
18 the other’s distress as their own distress. PD is strongly determined by the degree of attention
19 and sensitivity to the social signals expressed by others: someone highly self-centered may have
20 no contagion to prevent in the first place. Hence, the items of the PD subscale partly capture the
21 lack of SOD and partly capture the reactivity to others (e.g., “when I see someone who badly
22 needs help in an emergency, I go to pieces”). It is worth being noted that the inclusion of the PD
23 subscale to assess general empathy adds an unfitted load to the initial construct because well-

functioning SOD is theorized as a prerequisite of human empathy specifically (Decety & Jackson, 2004). Hence, the PD subscale should penalize the general empathy score rather than increase it. Unbeknownst to most researchers, the Empathic concern (EC) subscale also partly captures SOD: while it is defined as “the degree to which the respondent experiences feelings of warmth, compassion, and concern for the observed individual”, EC mechanistically results from functioning SOD since the items survey non-isomorphic emotions, that is, emotions that are distinct from those experienced by the empathized other person (e.g., “When I see someone being taken advantage of, I feel kind of protective toward them”). Such assessment of non-isomorphic emotions implies that the respondent tends not to mirror the other person's emotions, but instead, through a clear identification of their own and the other person's emotions, reacts in a complementary way, indicating a good understanding and self-regulation of the other person's emotions. Hence, EC partly captures the presence of SOD and partly captures the reactivity to others. Building on these observations, we have created a more specific measure of SOD with the PD and EC subscales of the IRI with the principle of cognitive subtraction. Specifically, if one subtracts the PD score from the EC score, it would capture SOD in both EC and PD contexts but with reactivity to others canceling each other (formula: $EC [Reactivity\ to\ others + presence\ of\ SOD] - PD [Reactivity\ to\ others + absence\ of\ SOD] = 2 \times presence\ of\ SOD$). Our first prediction was therefore that this SOD score would better differentiate ASPD individuals from healthy adults than the classic IRI subscales and total score.

Assessment of SOD via the level-2 Visual Perspective-Taking task

Many behavioral measures of mentalizing (or Theory of Mind) also partly capture SOD as long as it requires the participant to understand another person's thinking, which is conflicting with the participant's own thinking. Poor performance on these tasks is typically illustrated by

the *egocentric interference* (or bias or intrusion), that is, diminished performance caused by erroneously using information held from the egocentric perspective (e.g., self-projection) or struggling to prevent the use or influence of egocentric information in understanding others. For instance, Newbury-Helps et al. (2017) found higher egocentric interferences among 54 offenders diagnosed with ASPD in comparison to offenders without ASPD on the Director visual perspective-taking task (for a description of the task, see Dumontheil et al., 2010). However, in this task and all the other tasks measuring the egocentric interference, it is unclear whether this poorer performance is caused by a high self-focus (i.e., the tendency to prioritize information pertaining to the egocentric perspective over information related to other people) or by poor SOD. Thus, in their study, it cannot be excluded that the ASPD individuals were so self-focused that even with preserved SOD abilities, they struggled to not rely on their egocentric perspective and thus show a strong egocentric interference. Along the same line, some perspective-taking tasks include trials where the participant must adopt their egocentric perspective and report their own thinking, which is conflicting with the other person's thinking (for a review, see Quesque & Rossetti, 2020). Poor performance on these trials is typically illustrative of an *altercentric interference* (or bias or intrusion), that is, a diminished performance caused by erroneously using information inferred from the other person's perspective (i.e., altercentric information) or struggling to prevent the use or influence of altercentric information in understanding oneself. Hence, individuals showing an absence of altercentric interference could reflect either well-functioning SOD (interference is prevented) and/or insufficient prioritization of altercentric information, which includes a lack of attention, sensitivity, and reactivity to others (Lamm et al., 2016). Importantly, misunderstanding that the altercentric interference partly captures both SOD and the extent to which one prioritizes egocentric or altercentric information has led researchers

1 to arbitrarily interpret altercentric interference as good performance (sensitivity/concern for
2 others, high prioritization of others) and bad performance (poor SOD). Alternatively, an absence
3 of altercentric interference has been interpreted as impaired spontaneous perspective taking when
4 it could have been caused by well-functioning SOD (e.g., Drayton et al., 2018). To date, there
5 has been no specific measure of SOD in ASPD. However, we can exploit the fact that both the
6 egocentric and altercentric interferences capture (the lack of) SOD but also capture prioritization
7 of egocentric and altercentric information in opposite directions. Hence, combining (by summing
8 or averaging) the egocentric and altercentric interferences offers a specific measure of SOD since
9 the self-other prioritization elements cancel each other (formula: egocentric interference [lack of
10 SOD + egocentric prioritization] + altercentric interference [lack of SOD - egocentric
11 prioritization] = 2 times lack of SOD). Only a few perspective-taking paradigms include trials
12 requiring to adopt the other person's perspective but also the egocentric perspective and thus
13 allow to compute a SOD score. In this study, we used a level-2 visual perspective-taking task
14 (VPT; Surtees et al., 2016) which offers a performance-based measurement of perspective-taking
15 by evaluating an individual's ability to indicate *how* (relative to *what* for level-1 VPT; Flavell,
16 1977) another character sees an object when he or she is standing at the opposite end of a scene.
17 This task enables one to decompose between two dimensions: (1) the Self-other distinction
18 (SOD) dimension, which is the ability to distinguish between the self- and the other person's
19 perspective, and (2) the Self-other priority (SOP) dimension, which reflects the attentional
20 priority one mobilizes for information pertaining from the self-perspective in comparison to the
21 information pertaining from the other-perspective. This latter dimension allows to point out and
22 disentangle between the attentional biases for the self-versus another social agent (Bukowski,
23 2014; Bukowski & Samson, 2017), while the SOD dimension enables to capture the ability to

1 detect and resolve a conflict of viewpoints by inhibiting the self-perspective to adopt the other
2 person's perspective. This task offers significant advantages in that the low-level and implicit
3 processes that are targeted through a behavioral-based format cannot receive influences from
4 biases such as social desirability (as compared to self-report scales) or conscious motivation, and
5 allows an objective and sufficiently demanding assessment of both dimensions. Indeed,
6 participants are confronted with several trials in which the adoption and eventual selection of a
7 viewpoint are constrained by time pressure and require cognitive efforts (Surtees et al., 2016).
8 Our first prediction was therefore that this VPT SOD score would better differentiate ASPD
9 individuals from healthy adults than the classic egocentric and altercentric interferences scores
10 and the self-other priority (SOP dimension) score.

11 **Prediction 2: Objective measures of SOD better differentiate ASPDs from controls than**
12 **self-report measures**

13 From our hypothesis - H2 - that the failure to use the appropriate format to capture socio-
14 cognitive deficits in ASPD partly explains inconsistent results, we predicted that objective
15 performance-based measurement, as compared to self-report measurement, will better
16 differentiate ASPD individuals from healthy adults. Specifically, we predict that, while targeting
17 the right facet of social cognition is a necessary condition to detect a socio-cognitive impairment,
18 adopting the right format of assessment increases the capacity of the SOD measure to
19 differentiate ASPD from healthy individuals.

20 To recap, the exact cause of social inadaptation for ASPD is still unknown and calls for a
21 disentanglement between the distinct facets of social cognition, measures targeting lower-level
22 dimensions rather than broader concepts prone to confusion across studies, and the use of
23 objective and sufficiently demanding behavioral tasks rather than self-reports. In this study, we

tested ASPD and healthy adults with the IRI and a level-2 VPT task in order to disentangle two often confounded dimensions of social cognition, SOD and SOP, and to compare the respective explanatory power of a widely used self-report measure against a demanding performance-based perspective-taking task.

Materials and Method

Transparency and Openness

We report how we determined all data exclusions, all manipulations, and all measures in the study, and we follow JARS (Kazak, 2018). All data, analysis code, and research materials are available at <https://osf.io/4h256/>. Data were analyzed using IBM SPSS Statistics version 27.0 (IMB Corp., Amonk, NY, USA). This study's design and its analysis were not pre-registered.

Sample and Procedure

Sample is composed of 40 male participants, divided into two groups: 21 ASPD forensic inpatients (ASPD group) and 19 participants from the community (control group). A total of 73 participants were initially recruited, however, 10 participants were excluded because of missing data due to refusal or deterioration of mental state. All remaining participants with a VPT accuracy score below 66% in the congruent conditions were considered outliers (i.e. null hypothesis of random answer could not be excluded with 95% confidence; $N = 23$). The inclusion criterion for both groups was having French as their first language to understand instructions and questionnaires. The inclusion criteria for the ASPD group was to exhibit a stable mental state, assessed by their referring psychologist. ASPD forensic inpatients were hospitalized in a High-Risk Security Forensic Hospital "Les Marronniers" as they were recognized as Not Guilty for Reason of Insanity (Loi relative à l'internement des personnes,

2014). ASPD inpatients were individually assessed with the Mini International Neuropsychiatric Interview (MINI; Sheehan, 1998) and the Structured Clinical Interview for DSM-IV Axis II Disorders (SCID-II, First et al., 1995) at least one month after their admission to the High-Risk Security Forensic Hospital by ward psychologists trained in the administration of the instruments. Evaluations were carried out at Admission Services where the focus is on evaluating patients in a general way with no specific therapeutic aim. Forensic inpatients principally exhibited Axis I disorder (61.90%), such as Addictive Disorder (52.40%), and Axis I-II comorbidity (71.40%). Details of the prevalence of each disorder are reported in Table 1.

Table 1.

Prevalence of Axis-I and Axis-II disorders in our clinical sample.

	Axis-I					Axis-II			
Disorder	Psychotic	Mood	Addiction	Antisoc.	Narcissis.	Border.	OC	Parano.	Schizotyp.
N	2	5	11	21	6	11	2	2	1
%	9.5	23.5	52.4	100	28.6	52.4	9.5	9.5	4.5

Notes. Antisoc = antisocial; Narcissis = narcissistic; Border = borderline; OC = obsessive-compulsive; Parano = paranoid; Schizotyp = schizotypal. No anxiety, schizoid, histrionic, and dependent personality disorder were reported.

The overall mean duration of hospitalization for the ASPD group was 5.84 ($SD = 7.08$) years. In terms of offense history, ASPD forensic inpatients mainly committed current non-sexual non-violent offenses (76.2%), sexual offenses (52.40%), including rape (42.90%) and current violent offenses (52.40%), such as assault (28.60%).

Only forensic inpatients who matched the inclusion criteria were included in the ASPD sample by psychologists from the Forensic Hospital Care Units. Patients were presented with an information letter and a consent form, following the Declaration of Helsinki and the General Data Protection Regulation (GDPR; Regulation EU 2016/679). Participants from the control sample were solicited through social networks (Facebook, Instagram, etc.) and did not receive financial compensation. Both groups were tested in their respective place of residence and accompanied by an experimenter. This study was not preregistered.

Compliance with Ethical Standards

The ethics committee of UMonS granted approval to conduct this study (reference MEMANTEMO: CE/DV/EA/2015). The procedure was also conducted in accordance with the 1964 Declaration of Helsinki and its subsequent amendments or comparable ethical standards.

Instruments

Clinical Assessment

The Mini International Neuropsychiatric Interview (MINI; Sheehan et al., 1998) is a structured interview for adults used essentially to diagnose Axis I disorders according to the DSM-IV (American Psychiatric Association, 1994).

The Structured Clinical Interview for DSM-IV Axis-II Disorders (SCID-II, First et al., 1995) assesses the presence of 12 personality disorders according to the axis II of DSM-IV criteria based on 119 self-reported items along with a semi-structured interview to further explore the positive symptoms reported in the questionnaire (0 = “Absent or false”; 1 = “Subthreshold or questionable”; 2 = “Present or true”). Cluster A’s disorders refer to paranoid, schizoid, and schizotypal personalities, cluster B’s refer to antisocial, narcissistic, histrionic, and borderline personalities, and cluster C’s refer to avoidant, dependent and obsessive-compulsive

personalities. We used the French-validated version of SCID-II (Bouvard et al., 1999). The intraclass correlation coefficient for dimensional judgment ranged from .90-.98. Internal consistency, ranging from .71-.94, was satisfactory. All Cohen's Kappa coefficients for categorical diagnosis were $\geq .74$. These psychometric results suggest that both the interrater and internal consistency reliability of SCID-II are adequate (Maffei et al., 1997). Within the forensic inpatients sample, previous research highlighted that Cohen's Ks computed were $\geq .81$ for Clusters disorders (Vicenzutto et al., 2018).

Socio-cognitive Assessment

The Interpersonal Reactivity Index (IRI; Davis, 1983) is a self-reported questionnaire including 28 items with a 7-point Likert scale (from 1 = "does not describe me" to 7 = "perfectly describes me") measuring general empathy based on 4 distinct subscales, assessed by 7 items each; Perspective-taking and Empathic concern for cognitive empathy, and Fantasy and Personal distress for affective empathy. All standardized alpha coefficients vary between .68 and .79. In this study, we used Gilet et al. (2013)'s validated French version.

Subscales scores are calculated by computing the response to the related items¹. In line with our previous argument, we extracted a measure of SOD by calculating an index based on the Personal distress and Empathic concern scores of participants. We deduced the score on the PD subscale from the score on the EC subscale, both respectively capturing failed and successful self-other distinction in participants (see Table 1 of Supplemental Material).

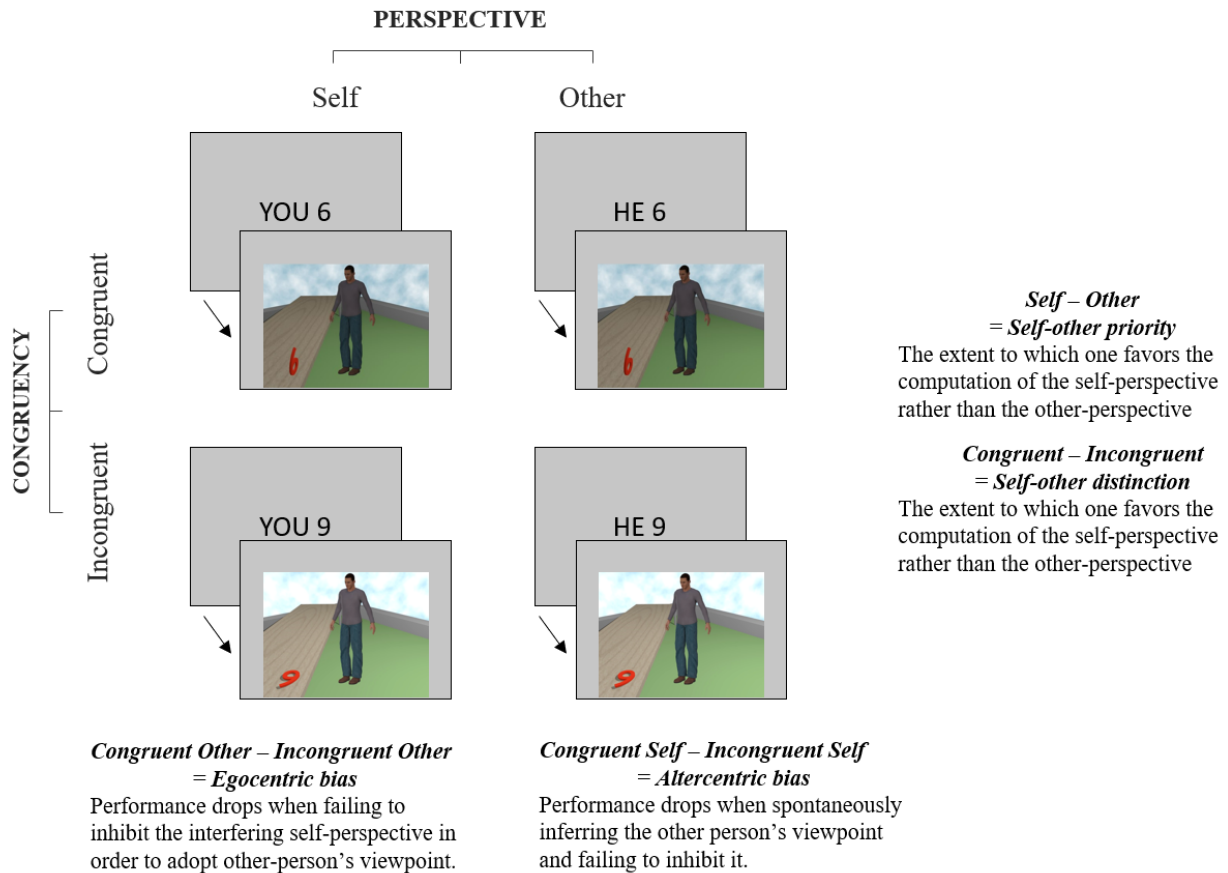
The Level-2 Visual Perspective-Taking task (VPT, Surtees et al., 2016) is a performance-based task that assesses the ability of a participant to infer how an object may look different to someone else. The participant is in front of a computer screen displaying a scene in which a

¹ One item was not completed by an ASPD participant (item 28). The missing value was replaced by the average score of the 6 other items relating to the perspective-taking subscale to compute the subscale and total scores.

human avatar stands next to a table, facing the participant. On this table is placed the number 6 or 9. This number can be laid down on the table, or be in a standing position. Depending on the trial's instruction, the participant is either asked to adopt his own point of view, or the avatar's (the "Perspective" condition in Figure 1). The position of the number changes pseudo-randomly to have 50% of congruent and 50% of incongruent trials (similar vs. different perspectives between the self and the avatar's; the "Congruency" condition in Figure 1). Each trial begins with a fixation cross displayed on the screen followed by the instruction indicating "YOU" or "HE/SHE" and the number "6" or "9". Next, the instruction disappears and the scene with the avatar looking at the number appears (see Figure 1). Participant has to press as fast as possible with the "C" key if the instruction was correct (i.e., indicating the right number according to the perspective the participant has to adopt), or "N" if it was incorrect. The task is composed of 24 training trials with feedback, followed by 96 experimental trials divided into two sequences of 5 minutes each. Participants' performance is assessed based on accuracy and response time. Four indexes were extracted based on participants' performance on each trial's condition.

Figure 1.

1 *Experimental design of the Level-2 VPT*



2

3 *Note.* Index calculation formulas for accuracy rates.

4

5 In line with previous studies using the Level 2 VPT task (Surtees et al., 2016), only

6 median response times (RT) from matching trials (i.e., "yes" response trials) were calculated, as

7 "no" responses are considered as fillers. Accuracy rates were calculated for all trials. Based on

8 scores obtained for both median RTs and accuracy rates, four indexes were calculated to

9 investigate Self-other priority and Self-other distinction dimensions of mentalizing:

10 The *SOP index* is used as a measure of *Self-other priority* (SOP), i.e., the extent to which

11 one favors the computation of the self-perspective rather than the other-perspective. It is

12 calculated by computing the difference in performance between self-perspective trials and other-

perspective trials ($RT_{\text{other}} - RT_{\text{self}}$; $Acc_{\text{self}} - Acc_{\text{other}}$). A positive index indicates a prioritization of the self-perspective compared to the other-perspective. In terms of response time, it indicates that the participant necessitates either a longer period to adopt the perspective of the character as compared to the adoption of the self-perspective, or, in terms of accuracy rates, commits fewer errors when having to adopt the self's perspective as compared to the character's perspective. Conversely, a negative index indicates a prioritization of the other-perspective.

The *SOD index* is used as a measure of *Self-other distinction* (SOD), i.e., the ability to detect and inhibit an interfering perspective. It is calculated by computing the difference between the performance in congruent trials and incongruent trials ($RT_{\text{inc}} - RT_{\text{con}}$; $Acc_{\text{con}} - Acc_{\text{inc}}$). A positive index indicates a tendency to receive interferences from an irrelevant perspective that can be translated by longer response latencies or an increase of errors when being presented with conflicting perspectives, whereas a negative score indicates a good ability to inhibit an interfering perspective.

The *egocentric bias* reflects the tendency to project the self-perspective onto the other when perspectives differ and thus relates to the common definition of egocentrism. It is calculated by computing the difference between the performance of congruent other-condition trials and incongruent other-condition trials ($RT_{\text{other_inc}} - RT_{\text{other_con}}$; $Acc_{\text{other_con}} - Acc_{\text{other_inc}}$). A higher index indicates a higher extent of interferences received from the self-perspective.

The *altercentric bias* reflects the tendency to receive interferences from the other-perspective when adopting the self-perspective. It is calculated by computing the difference between the performance of congruent self-condition trials and incongruent self-condition trials ($RT_{\text{self_inc}} - RT_{\text{self_con}}$; $Acc_{\text{self_con}} - Acc_{\text{self_inc}}$). A higher index indicates a higher extent of interferences received from the other-perspective.

Results

Demographic differences

ASPD group and control group significantly differed in terms of age ($t(38) = 3.003, p = .005$) and mean education level ($t(38) = -10.711, p < .001$), with a mean age of 43.00 ($SD = 11.40$) years old for the ASPD group and a mean age of 32.32 ($SD = 11.03$) for the control group. The mean education level for the ASPD group was 7.81 ($SD = 2.62$) years, and 14.53 ($SD = 1.12$) years for the control group. To ensure the absence of confounding effects on our results, bivariate correlations were conducted on separated groups between both age and education level and SOP and SOD indexes (accuracy rates and median RTs). Two regression models were also performed with either age or education level as predictor, and with the group as a control variable to predict SOP and SOD indexes. Age was significantly associated with the median RT of the SOP index ($r = .661, p = .002$) in the control group, but no other correlations reached significance across groups ($ps > .074$), and no effect of age or education level was reported in regression models when controlling for group (all $Bs < .005$, all $ps > .275$).

Considering the significant differences in age and education between the groups, the high range of RTs (in contrast to accuracy that ranges from 66% to 100%), and how RTs are affected by age and education, each RT index was corrected by the overall speed level of the participant (e.g., $SOD\ RT\ index = (RT_{inc} - RT_{con})/RT_{mean}$). Such a procedure enables to reduce the impact of group differences in domain-general cognitive functioning (caused by age, education, medication, or else) in order to capture group differences that are more likely to be specifically socio-cognitive in nature.

1 **H.1. Which socio-cognitive facet is impaired in ASPD?**

2 Our first main hypothesis was to find lower social cognition scores for the ASPD group
 3 as compared to the control group, specifically on the SOD dimension (see Table 2). Multiple
 4 student t-tests for independent samples were performed on each socio-cognitive measure, namely
 5 the 4 VPT indexes, both on median response times² and accuracy rates, as well as on the IRI's
 6 total and subscales scores (see Table 3 of Supplemental Material). Welsh's correction was
 7 performed for accuracy rates due to significant heterogeneity of variances. Decomposed
 8 predictions are reported below.

9 First, we posited to find a specific deficit in the SOD dimension for the ASPD group. We
 10 expected to find significant differences in the IRI's and the VPT's SOD index, but not in the
 11 SOP index. Results revealed a significant difference between both groups on the VPT's SOD
 12 index, with more errors in the ASPD group due to confusion between the self and the other's
 13 perspectives ($t(38) = -3.275, p = .003; M_{\text{ASPD}} = 19.07\%, SD = 20.13\%$ and $M_{\text{control}} = 4.26\%, SD =$
 14 4.69%). In other words, when facing conflicting viewpoints (in comparison to congruent
 15 viewpoints), the drop-in accuracy at judging visual perspectives was significantly higher for the
 16 ASPD group than for the control group, with an increase of 19 percentage points in errors in the
 17 ASPD group compared to an increase of 4 percentage points in the control group. In addition, a
 18 significant difference between the two groups was observed in the IRI's SOD scores, indicating a
 19 lower SOD ability in ASPD participants than in controls ($t(38) = 2.511, p = .016; M_{\text{ASPD}} = 7.00,$
 20 $SD = 7.32$ and $M_{\text{control}} = 12.95, SD = 7.66$). No significant difference was observed on the SOP
 21 index, suggesting that ASPDs did not tend to respond more easily when having to adopt either
 22 perspectives (the self or the other's) as compared to controls.

² One patient was excluded from the median RTs analyses due to an encoding system failure.

Second, we predicted finding group differences on the egocentric and altercentric bias indexes, as both measures are posited to partly tap into SOD ability. Results partly confirmed our predictions, with a significant difference between groups observed on the egocentric bias index. This difference suggests that the ASPD participants tended to commit more mistakes due to interferences stemming from their own perspective ($t(38) = -2.762, p = .012; M_{\text{ASPD}} = 21.05\%, SD = 30.89\%; M_{\text{control}} = 2.11\%, SD = 5.49\%$) than the control participants. In other words, when facing conflicting viewpoints and having to adopt the other's perspective (in comparison to non-conflicting viewpoints trials requiring to adopt the other person's perspective), the drop-in accuracy was significantly higher for the ASPD group than for the control group. Although no difference was found in median response times, numerical differences observed in each group suggested a greater difference in response latency when facing conflicting viewpoints while having to adopt the other's perspective (compared to congruent trials with an instruction to adopt the other's perspective) in ASPD group than in the control group ($M_{\text{ASPD}} = 0.13, SD = 0.19; M_{\text{control}} = 0.21, SD = 0.20$).

Regarding the altercentric bias, the ASPD participants numerically made more mistakes due to interferences stemming from spontaneously tracking the other person's perspective than the control participants, but the difference did not reach statistical significance, ($t(38) = -1.663, p = .109; M_{\text{ASPD}} = 17.10\%, SD = 27.87; M_{\text{control}} = 6.42\%, SD = 8.95\%$). The same pattern was observed for RTs, ($t(37) = -0.880, p = .384; M_{\text{ASPD}} = 0.103, SD = 0.15; M_{\text{control}} = 0.02, SD = 0.40$).

Finally, we predicted a significant group difference in the IRI's EC and PD scores, as these measures are posited to partly tap into SOD. On the other hand, neither the IRI's composite score nor the IRI's PT score was expected to account for the SOD process, therefore no

1 difference was expected. Results revealed a significant difference between the ASPD group and
2 control group on the EC scores, with controls reporting feeling more frequently sympathy and
3 compassion for others ($t(38) = 2.374, p = .023$) than ASPDs ($M_{\text{ASPD}} = 32.24, SD = 5.16; M_{\text{control}}$
4 $= 36.11, SD = 5.13$). Although no significant difference was found on the PD subscale, mean
5 scores suggested that the ASPD group reported feeling more overwhelmed when confronted with
6 the others' negative affective states ($M_{\text{PD}} = 25.24, SD = 5.67$) than the control group ($M_{\text{PD}} =$
7 $23.16, SD = 7.21$). Finally, no significant group difference was found on the PT subscale score
8 ($M_{\text{ASPD}} = 30.14, SD = 9.38; M_{\text{control}} = 33.47, SD = 6.17$) nor the IRI composite score ($M_{\text{ASPD}} =$
9 $121.10, SD = 15.29; M_{\text{control}} = 122.68, SD = 15.21$) between the two groups, indicating that the
10 reported motivation to adopt the perspective of the others and the general sensitivity to others did
11 not differ in the ASPD compared to the control group.

12 **H.2. Which socio-cognitive assessment method best characterizes impairment in** 13 **ASPDs?**

14 Our second main hypothesis was to find clearer evidence of group differences between
15 ASPD and control groups using performance-based measures (VPT's indexes) compared to self-
16 reports (IRI's composite and subscales scores). Bivariate correlations were performed to test this
17 assumption, along with logistic regressions to predict group membership. Decomposed
18 predictions are reported below.

19 We first posited that the IRI's SOD score would be significantly associated with the
20 VPT's SOD index, as both measures are assumed to capture the SOD dimension. Second, a
21 similar association between IRI's PT and VPT's SOP index was expected, both measures being
22 assumed to reflect the SOP dimension. A negative significant correlation between the IRI's SOD
23 and the VPT's SOD index was found, suggesting that a higher difference between the EC and

PD scores would predict fewer errors during visual perspective-taking due to confusion between the self and the other's viewpoints ($r(39) = -0.355, p = .025$). In other words, participants who showed a stronger contrast between their EC (functioning SOD) and PD (dysfunctioning SOD) scores tended to exhibit a smaller drop in accuracy at judging visual perspectives when facing conflicting viewpoints (in comparison to congruent viewpoints). This observation supports our assumption that both measures reflect the ability to distinguish between the self and the others' mental states. In the same vein, a negative correlation was found between the IRI's PT and the VPT's SOP median response time ($r(39) = -0.401, p = .011$), indicating that participants who scored higher on the PT subscale tended to respond faster when they were required to adopt the other's perspective compared to their own in the VPT task. This result replicates the significant associations previously found in studies conducted on samples of students and narcissistic participants while controlling for the level of narcissism (Bukowski & Samson, 2017; Bukowski et al., 2022). This finding supports our assumption that both measures reflect one's motivation and sensitivity toward the other.

Table 2.

Summary of results of our hypotheses and predictions

Why inconsistent results in characterizing ASPD socio-cognitive difficulties?		H1: Failure to specifically identify the impaired facet of social cognition		
		SOD not measured	SOD non-specific	SOD specific
H2. Failure to use the appropriate format to capture socio-cognitive deficits	Self-report measure (IRI)	IRI PT	IRI EC	IRI SOD
		$p = .248$	$p = .023$	$p = .016$
		$d = 0.420$	$d = 0.752$	$d = 0.794$
			IRI PD	
			$p = .389$	
			$d = 0.321$	

Objective performance task (VPT)	VPT SOP $p = .149$ $d = 0.462$	VPT egocentric bias	
		$p = .012$	VPT SOD $p = .003$ $d = 1.013$
		$d = 0.854$	
		VPT altercentric bias	
		$p = .109$	
		$d = 0.516$	
Best measure: conclusions	VPT>IRI	SOD not measured < SOD non-specific < SOD specific	

Second, we predicted that the use of performance-based measures would be better suited to characterize social cognition impairments in ASPD, and so that group membership would be better predicted by the VPT's SOD index than the IRI's SOD. Both variables having already proved their relevance to significantly differentiate between each group, we entered IRI and VPT'SOD index in a two-steps binary logistic regression model to predict group membership. IRI's SOD index was introduced as the unique predictor in the first step of the model, and VPT's SOD index was then added in the second step of the model. We used the Likelihood Ratio Test at each step to assess whether adding VPT's SOD index significantly improved the prediction of group membership, thereby indicating significant incremental information provided by a score based on a performance format versus a self-report format. In line with our previous results, a significant contribution of IRI's SOD to predict group membership ($\chi^2_{\text{step 1}}(1, N = 40) = 6.089, p = .014$, Nagelkerke's $R^2 = .188$) was found when entered as the unique predictor. We then entered VPT's SOD index, which indicated a significant improvement in the prediction of group membership ($\chi^2_{\text{step 2}}(1, N = 40) = 6.380, p = .012$; Nagelkerke's $R^2 = .357$). Conducted the other way around, adding the IRI's SOD to the VPT's SOD index did not significantly improve group prediction ($\chi^2_{\text{step 2}}(1, N = 40) = 2.798, p = .094$), hence confirming the significant added value of

1 using a SOD performance index compared with the use of a SOD score based on a self-report
2 assessment.

4 **Discussion**

5 Previous studies have attempted to identify and understand which socio-cognitive facets
6 may be impaired in individuals with ASPD, however, inconsistent results have emerged despite
7 the growing body of research (Campos et al., 2022; Chang et al., 2021; Song et al., 2023). In this
8 study, we sought to address this issue by investigating a more specific socio-cognitive
9 dimension, the self-other distinction (SOD) ability, and by providing an adequate methodology
10 to capture this deficit. To this end, we administered two distinct formats of socio-cognitive
11 measures to our participants, namely the IRI and the level-2 VPT task. Results evidenced a
12 significant difference between the two groups on the SOD dimension, with more difficulties in
13 the ASPD group. Furthermore, we compared the use of a self-report format with a performance-
14 based format to capture SOD impairment and found a superior contribution of performance-
15 based scores in predicting group membership. These results, their interpretations, implications,
16 and limits are discussed below.

17 **Self-Other Distinction (SOD) as a core deficit in ASPD**

18 We hypothesized the existence of an impairment in the ability to distinguish between
19 self- and others' mental states in people with antisocial personality disorder. We found that, in
20 comparison to a group of healthy control participants, ASPD participants were significantly and
21 specifically more inaccurate at judging someone's visual perspective (their own or those of
22 another person) when another concomitant and differing perspective was present. Our results
23 supported this hypothesis, indicating that people with antisocial personality disorder struggle to

distinguish between information coming from both mental states. Distinguishing between the self and other is a facet of social cognition that is fundamental to navigating the interpersonal world, which is found dysfunctional in antisocial personality disorder, as evidenced by the high proportion of criminality in this population (Black, 2010). Differentiating between our own mental state and those of others involves the ability to identify what stems from the self, such as emotional activation or cognitions, and what stems from others, in order to avoid confusion and preserve the integrity of the self. The inability or lack of attempt to distinguish between both mental states often manifests in excessive reliance on rapid, automatic and biased mentalization rather than exerting cognitive control that prevents confusion between the different information coming from both sides (Luyten et al., 2021). Consequently, the confusion between these sources of information may lead to an increased vulnerability to emotional contagion, which can culminate in emotional outbursts often observed in ASPD individuals (Martin et al., 2019; Nalbant, 2022; Velotti et al., 2019). In that sense, the combination of a weakened sense of self, impaired inhibitory affective control and aggression can be interpreted as the manifestation of a defective SOD. In light of Eddy (2018)'s assumption, the externalized behaviors displayed by ASPDs, such as aggression for example, may be read as a coping mechanism for the tenuous boundaries between self and others. The adoption of antagonistic behaviors can then be displayed in an attempt to separate the individual from his or her counterpart and overcome excessive emotional contagion responsible for personal distress, to recover proper distinction between the self and the other (Eddy, 2022).

Self-Other Priority (SOP), the hidden twin brother of SOD

In this study, we examined SOP along with SOD as two distinct dimensions underlying mentalizing performance. The degree to which the ASPD participants performed better at

judging from their own visual perspective than the perspective of the other person was not significantly higher than in the group of healthy control participants. Hence, our results do not support the hypothesis that the ASPD individuals prioritize their egocentric perspective to a particularly high extent. Self-other priority is not an ability per se but an indication of the extent to which one is more efficient at processing self-information than information about other people. Examples of confounds between SOD and SOP are frequent in ASPD research (e.g., Newbury-Helps et al., 2017). Based on our measure of SOP in the VPT task but also, indirectly, via the IRI total score and the IRI PT scores that both capture SOP as well, our results suggest that ASPD individuals are not particularly self-centered nor unmotivated to adopt the perspective of the others, and that previous studies' findings suggesting so should rather be interpreted as poor SOD (see Supplemental Materials for further discussion).

The adequacy of the socio-cognitive assessment format

We hypothesized that the inconsistent pattern of socio-cognitive deficits in ASPD research may be partially due to the inadequate format used to capture specific socio-cognitive deficits. Indeed, previous studies have mainly focused on self-report assessment (Jolliffe & Farrington, 2004; van Langen et al., 2014) despite the significant risk of biases present in the ASPD population (Cankaya et al., 2022; Paulhus & Williams, 2002; Velotti et al., 2019). In addition, when using performance-based measures, task designs have not always provided the conditions necessary to highlight specific impairments (e.g., Dolan & Fullam, 2004).

By comparing the specific contributions of each mentalizing assessment format (performance-based versus self-reported), we demonstrated the significantly higher contribution of using an objective measure of SOD over its self-report counterpart to differentiate ASPD versus healthy individuals. This result aligns with the premise that sufficiently demanding

assessment conditions are required to better identify and understand the socio-cognitive deficits in ASPD, but also on a broader scale.

In this study, we have mainly focused on the assessment of the SOP and SOD dimensions of social cognition. While we reported some important criticisms regarding self-reported evaluation in ASPD, we demonstrated that using a novel index of SOD based on IRI scores was relevant to assess this specific facet, as it showed to be significantly associated with the SOD index obtained with a behavioral task and differentiated the ASPDs from the control group. Further, we replicated the significant correlation between the SOP index and the IRI's PT subscale found in previous studies (Bukowski et al., 2022; Bukowski & Samson, 2017), hence confirming the ability of the VPT task to, first, precisely quantify the attentional priority allocated for self-information compared to information about others (SOP), and second, to specifically capture the ability to disentangle information coming from the self and from the others (SOD).

What is captured by the egocentric bias?

Our results demonstrated a significant difference on the egocentric bias between the control group and the ASPD group. We seek to clarify that the egocentric bias partly reflects the tendency to focus attention on the self-perspective, which is tied to the SOP score, and partly the ability to suppress the irrelevant points of view (here, the self-perspective), which is tied to the SOD score. Yet, because a group difference was found on the SOD score, the group difference found on the egocentric bias was most likely due to a failure to disentangle between the self and the other perspective (see Supplemental Materials for further discussion).

Limits and Future Perspectives

This study has several limitations. First, the VPT task is a low-level cognitive behavioral measure that enables a streamlined assessment of implicit and explicit processes. However, due to its simplified design of processing visual experiences, further investigation of a more affective side of SOD should be conducted. For instance, the visuo-tactile empathy task developed by Silani et al. (2013) is based on a two-dimensional paradigm that is similar to the one used in the VPT task, thus allowing to measure both SOP and SOD dimensions. In that task, participants are required to judge the level of agreeability or discomfort of the self or someone else's sensorial experience (pain or softness applied on the hand) while experiencing a congruent or incongruent sensitive stimulus. Confronting current findings on the cognitive computation of viewpoints, with the computation of affective experiences could lead to a better understanding of the extent of SOD impairment in ASPD individuals.

Secondly, our clinical sample included individuals presenting a common diagnosis of antisocial personality and other psychological disorders such as psychosis or borderline personality disorder, for instance. Due to this heterogeneity, we cannot fully ascertain that ASPD features were uniquely responsible for SOD impairment. However, despite the theoretical approach of categorical and dimensional conceptualization of mental health disorders, the reality of the investigated population lies in the presence of comorbidities (Oswald et al., 2017; Pham & Saloppé, 2010). Future studies should attempt to replicate the current findings in ASPD samples by recruiting participants from populations less likely to present multiple disorders, such as the carceral population, for instance. On the other hand, taking into account the potential effect of comorbid traits or psychopathologies may provide valuable insight into the predictive characteristics of SOD impairment. By doing so, it would enable to (1) identify clinical profiles

1 that are more prone to present SOD impairments and (2), to potentially investigate SOD as a
2 transdiagnostic component of disorders characterized by socio-cognitive deficits. Indeed, various
3 clinical signs displayed in clinical populations with distinct diagnoses may in fact reflect the
4 same underlying SOD impairment. For example, individuals with borderline personality are
5 often victims of self-injury, which may in fact reflect an attempt to regain a sense of agency and
6 cope with a reduced SOD (Eddy, 2022). Adopting a transdiagnostic approach offers a novel
7 understanding of the prominent heterogeneity and similarities found across psychopathologies,
8 and encourages the use of clinical applications that (at least partly) tackle the ability to handle
9 conflicting mental states. One example of such intervention is the mentalizing-based treatment
10 (MBT) developed by Bateman & Fonagy (2010) which focuses on stabilizing a sense of self, and
11 fostering the capacity to reinstate or maintain mentalizing when it is or is being lost. This
12 intervention is carried out in five stages, during which the patients are trained to examine their
13 thoughts and feelings while adopting an alternative perspective of a situation. By encouraging
14 the adoption of a distinct perspective, patients are led to compute their own perception and the
15 alternative perception, and to manage the ensuing conflict.

16 Third, in line with the present findings, future research investigating SOD ability or other
17 facets of social cognition in ASPD should strive to adopt performance-based assessment methods
18 to prevent the obtention of biased or inconclusive results due to inherent biases of self-reported
19 methods. Promising work is emerging in this direction (e.g., Krol & Bartz, 2022).

20 Fourth, some limitations can be pointed out regarding our sample characteristics. First,
21 our sample sizes were relatively small and only included male adults, which may raise concerns
22 regarding the representativity of the ASPD population. Second, our control sample failed to
23 match the sociodemographic profiles of the clinical group. Despite additional analyses, we

cannot exclude the influence of group differences in domain-general cognitive functioning (caused by age, education, medication, or else). These differences between groups may be partly due to the method used to recruit our control participants. Indeed, by using social media, we indirectly favored the accessibility of individuals who are familiar with computer technology and related tools. Beyond familiarity, other limitations can be stated with regard to this sampling method: participants had to have or have access to a computer and know how to use it, they had to be aware of the existence of this study (which refers to the management of publicity for this study), and they had to be interested in participating despite the absence of reward or compensation (which may trigger selection bias such as participants with an oriented interest). We encourage the testing of larger, mixed samples that would enable gender and comorbidities to be integrated into statistical models. Future studies in the field should further investigate SOD deficit as a transdiagnostic impairment characterizing clinical status, and systematic assessments of social cognition should be promoted by facilitating the use and accessibility of tests.

Conclusion

To conclude, our findings suggest that ASPD individuals' poor socio-cognitive functioning may stem specifically from impaired self-other distinction, a key mechanism in both social cognition and a predominant personality disorder theoretical framework. Our findings also offer some insights as to the methodological limitations that might partly explain the equivocal findings reported in previous ASPD studies. At last, our findings demonstrated the interest in clearly defining and teasing apart the facets of social cognition and using tools that are objectively aligned with their construct, and sufficiently stringent to capture real-life difficulties.

References

- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders (DSM-IV)*.
- Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The “Reading the Mind in the Eyes” Test Revised Version: A Study with Normal Adults, and Adults with Asperger Syndrome or High-functioning Autism. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, 42(2), 241–251.
<https://doi.org/10.1017/S0021963001006643>
- Bateman, A., Bolton, R., & Fonagy, P. (2013a). Antisocial Personality Disorder: A Mentalizing Framework. *FOCUS*, 11(2), 178–186. <https://doi.org/10.1176/appi.focus.11.2.178>
- Bateman, A., Bolton, R., & Fonagy, P. (2013b). Antisocial Personality Disorder: A Mentalizing Framework. *FOCUS*, 11(2), 178–186. <https://doi.org/10.1176/appi.focus.11.2.178>
- Bateman, A., & Fonagy, P. (2010). Mentalization based treatment for borderline personality disorder. *World Psychiatry*, 9(1), 11–15.
- Black, D. W. (2010). Antisocial personality disorder in incarcerated offenders: Psychiatric comorbidity and quality of life. *ANNALS OF CLINICAL PSYCHIATRY*.
- Bigot, A., Saloppé, X., Jean-Louis Nandrino, Pham, T., Bukowski, H., & Tiberi, L. A. (2023). *Self-Other Distinction in ASPD*. <https://doi.org/10.17605/OSF.IO/4H256>
- Bouvard, M., Fontaine-Buffer, M., Cungi, C., Adeleine, P., Chapoutier, C., Durafour, E., Bouchard, C., & Cottraux, J. (1999). Etude préliminaire d’un entretien structuré des troubles de la personnalité: Le SCID II. [Preliminary study of the structured diagnostic interview for personality disorders: The SCID II.]. *L’Encéphale: Revue de Psychiatrie Clinique Biologique et Thérapeutique*, 25, 416–421.

- 1 Bukowski, H. (2014). *What influences perspective taking ? A dynamic and multidimensional*
2 *approach* [UCL - Université Catholique de Louvain].
3 <https://dial.uclouvain.be/pr/boreal/object/boreal:151995>
- 4 Bukowski, H. (2019). Self-Knowledge. In V. Zeigler-Hill & T. K. Shackelford (Eds.),
5 *Encyclopedia of Personality and Individual Differences* (pp. 1–7). Springer International
6 Publishing. https://doi.org/10.1007/978-3-319-28099-8_2004-1
- 7 Bukowski, H., bigot, A., & Samson, D. (2022). *Narcissists are not more egocentric: Evidence*
8 *from a performance-based perspective-taking measure* [Preprint]. PsyArXiv.
9 <https://doi.org/10.31234/osf.io/rtyfm>
- 10 Bukowski, H., & Samson, D. (2017). New Insights into the Inter-Individual Variability in
11 Perspective Taking. *Vision*, 1(1), 8. <https://doi.org/10.3390/vision1010008>
- 12 Campos, C., Pasion, R., Azeredo, A., Ramião, E., Mazer, P., Macedo, I., & Barbosa, F. (2022).
13 Refining the link between psychopathy, antisocial behavior, and empathy: A meta-
14 analytical approach across different conceptual frameworks. *Clinical Psychology Review*,
15 94, 102145. <https://doi.org/10.1016/j.cpr.2022.102145>
- 16 Cankaya, H., Cakmak, S., Tamam, L., Namli, Z., Demirkol, M., & Karaytug, M. (2022). The
17 Relationship Between Suicidal Behavior and Metacognitive Characteristics in Male
18 Patients with Antisocial Personality Disorder. *Psychiatry and Behavioral Sciences*, 12(4),
19 188. <https://doi.org/10.5455/PBS.20220625091308>
- 20 Chang, S.-A. A., Tillem, S., Benson-Williams, C., & Baskin-Sommers, A. (2021). Cognitive
21 Empathy in Subtypes of Antisocial Individuals. *Frontiers in Psychiatry*, 12.
22 <https://www.frontiersin.org/articles/10.3389/fpsy.2021.677975>

- 1 Dang, J., King, K. M., & Inzlicht, M. (2020). Why Are Self-Report and Behavioral Measures
2 Weakly Correlated? *Trends in Cognitive Sciences*, 24(4), 267–269.
3 <https://doi.org/10.1016/j.tics.2020.01.007>
- 4 Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a
5 multidimensional approach. *Journal of Personality and Social Psychology*, 44, 113–126.
6 <https://doi.org/10.1037/0022-3514.44.1.113>
- 7 Decety, J., & Jackson, P. L. (2004). The Functional Architecture of Human Empathy. *Behavioral*
8 *and Cognitive Neuroscience Reviews*, 3(2), 71–100.
9 <https://doi.org/10.1177/1534582304267187>
- 10 Dimitrijević, A., Hanak, N., Altaras Dimitrijević, A., & Jolić Marjanović, Z. (2018). The
11 Mentalization Scale (MentS): A Self-Report Measure for the Assessment of Mentalizing
12 Capacity. *Journal of Personality Assessment*, 100(3), 268–280.
13 <https://doi.org/10.1080/00223891.2017.1310730>
- 14 Dolan, M., & Fullam, R. (2004). Theory of mind and mentalizing ability in antisocial personality
15 disorders with and without psychopathy. *Psychological Medicine*, 34(6), 1093–1102.
16 <https://doi.org/10.1017/S0033291704002028>
- 17 Drayton, L. A., Santos, L. R., & Baskin-Sommers, A. (2018). Psychopaths fail to automatically
18 take the perspective of others. *Proceedings of the National Academy of Sciences*,
19 115(13), 3302–3307. <https://doi.org/10.1073/pnas.1721903115>
- 20 Dumontheil, I., Apperly, I. A., & Blakemore, S.-J. (2010). Online usage of theory of mind
21 continues to develop in late adolescence. *Developmental Science*, 13(2), 331–338.
22 <https://doi.org/10.1111/j.1467-7687.2009.00888.x>

- 1 Eddy, C. M. (2018). Social cognition and self-other distinctions in neuropsychiatry: Insights
2 from schizophrenia and Tourette syndrome. *Progress in Neuro-Psychopharmacology and*
3 *Biological Psychiatry*, 82, 69–85. <https://doi.org/10.1016/j.pnpbp.2017.11.026>
- 4 Eddy, C. M. (2019). What Do You Have in Mind? Measures to Assess Mental State Reasoning
5 in Neuropsychiatric Populations. *Frontiers in Psychiatry*, 10.
6 <https://www.frontiersin.org/articles/10.3389/fpsyt.2019.00425>
- 7 Eddy, C. M. (2022). The Transdiagnostic Relevance of Self-Other Distinction to Psychiatry
8 Spans Emotional, Cognitive and Motor Domains. *Frontiers in Psychiatry*, 13, 797952.
9 <https://doi.org/10.3389/fpsyt.2022.797952>
- 10 First, M. B., Spitzer, R. L., Gibbon, M., Williams, J. B. W., Davies, M., Borus, J., Howes, M. J.,
11 Kane, J., Pope, H. G., & Rounsaville, B. (1995). The Structured Clinical Interview for
12 DSM-III-R Personality Disorders (SCID-II). Part II: Multi-Site Test-Retest Reliability
13 Study. *Journal of Personality Disorders*, 9(2), 92–104.
14 <https://doi.org/10.1521/pedi.1995.9.2.92>
- 15 Flavell, J. H. (1977). The development of knowledge about visual perception. *Nebraska*
16 *Symposium on Motivation. Nebraska Symposium on Motivation*, 25, 43–76.
- 17 Fonagy, P., & Luyten, P. (2009). A developmental, mentalization-based approach to the
18 understanding and treatment of borderline personality disorder. *Development and*
19 *Psychopathology*, 21(4), 1355–1381. <https://doi.org/10.1017/S0954579409990198>
- 20 Gilet, A.-L., Mella, N., Studer, J., Grün, D., & Labouvie-Vief, G. (2013). Assessing
21 dispositional empathy in adults: A French validation of the Interpersonal Reactivity Index
22 (IRI). *Canadian Journal of Behavioural Science / Revue Canadienne Des Sciences Du*
23 *Comportement*, 45, 42–48. <https://doi.org/10.1037/a0030425>

- 1 Hall, J. A., & Schwartz, R. (2019). Empathy present and future. *The Journal of Social*
2 *Psychology, 159*(3), 225–243. <https://doi.org/10.1080/00224545.2018.1477442>
- 3 Hogan, R. (1969). *Development of an empathy scale. 33*(3), 307–316.
4 <https://doi.org/10.1037/h0027580>
- 5 Johnson, B. N., Kivity, Y., Rosenstein, L. K., LeBreton, J. M., & Levy, K. N. (2022). The
6 association between mentalizing and psychopathology: A meta-analysis of the reading
7 the mind in the eyes task across psychiatric disorders. *Clinical Psychology: Science and*
8 *Practice, 29*, 423–439. <https://doi.org/10.1037/cps0000105>
- 9 Jolliffe, D., & Farrington, D. P. (2004). Empathy and offending: A systematic review and meta-
10 analysis. *Aggression and Violent Behavior, 9*(5), 441–476.
11 <https://doi.org/10.1016/j.avb.2003.03.001>
- 12 Kazak, A. E. (2018). Editorial: Journal article reporting standards. *American Psychologist, 73*(1),
13 1–2. <https://doi.org/10.1037/amp0000263>
- 14 Kittel, A. F. D., Olderbak, S., & Wilhelm, O. (2022). Sty in the Mind’s Eye: A Meta-Analytic
15 Investigation of the Nomological Network and Internal Consistency of the “Reading the
16 Mind in the Eyes” Test. *Assessment, 29*(5), 872–895.
17 <https://doi.org/10.1177/1073191121996469>
- 18 Krol, S. A., & Bartz, J. A. (2022). The self and empathy: Lacking a clear and stable sense of self
19 undermines empathy and helping behavior. *Emotion, 22*(7), 1554–1571.
20 <https://doi.org/10.1037/emo0000943>
- 21 Lamm, C., Bukowski, H., & Silani, G. (2016). From shared to distinct self–other representations
22 in empathy: Evidence from neurotypical function and socio-cognitive disorders.

- 1 *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1686),
2 20150083. <https://doi.org/10.1098/rstb.2015.0083>
- 3 Lavallee, A., Gandolphe, M.-C., Saloppé, X., Ott, L., Pham, T., & Nandrino, J.-L. (2020).
4 Characterisation of self-defining memories in criminals with antisocial personality
5 disorder. *Memory*, 28(9), 1123–1135. <https://doi.org/10.1080/09658211.2020.1818785>
- 6 Lilienfeld, S. O., & Strother, A. N. (2020). Psychological measurement and the replication crisis:
7 Four sacred cows. *Canadian Psychology / Psychologie Canadienne*, 61, 281–288.
8 <https://doi.org/10.1037/cap0000236>
- 9 Luyten, P., De Meulemeester, C., & Fonagy, P. (2021). The Self–Other Distinction in
10 Psychopathology: Recent Developments from a Mentalizing Perspective. In M. Gilead &
11 K. N. Ochsner (Eds.), *The Neural Basis of Mentalizing* (pp. 659–680). Springer
12 International Publishing. https://doi.org/10.1007/978-3-030-51890-5_34
- 13 Maffei, C., Fossati, A., Agostoni, I., Barraco, A., Bagnato, M., Deborah, D., Namia, C., Novella,
14 L., & Petrachi, M. (1997). Interrater Reliability and Internal Consistency of the
15 Structured Clinical Interview for DSM-IV Axis II Personality Disorders (SCID-II),
16 Version 2.0. *Journal of Personality Disorders*, 11(3), 279–284.
17 <https://doi.org/10.1521/pedi.1997.11.3.279>
- 18 Malterer, M. B., Glass, S. J., & Newman, J. P. (2008). Psychopathy and trait emotional
19 intelligence. *Personality and Individual Differences*, 44(3), 735–745.
20 <https://doi.org/10.1016/j.paid.2007.10.007>
- 21 Marsden, J., Glazebrook, C., Tully, R., & Völlm, B. (2019). Do adult males with antisocial
22 personality disorder (with and without co-morbid psychopathy) have deficits in emotion

- processing and empathy? A systematic review. *Aggression and Violent Behavior*, 48, 197–217. <https://doi.org/10.1016/j.avb.2019.08.009>
- Martin, S., Zabala, C., Del-Monte, J., Graziani, P., Aizpurua, E., Barry, T. J., & Ricarte, J. (2019). Examining the relationships between impulsivity, aggression, and recidivism for prisoners with antisocial personality disorder. *Aggression and Violent Behavior*, 49, 101314. <https://doi.org/10.1016/j.avb.2019.07.009>
- Mehrabian, A., & Epstein, N. (1972). A measure of emotional empathy1. *Journal of Personality*, 40(4), 525–543. <https://doi.org/10.1111/j.1467-6494.1972.tb00078.x>
- Miller, P. A., & Eisenberg, N. (1988). The relation of empathy to aggressive and externalizing/antisocial behavior. *Psychological Bulletin*, 103, 324–344. <https://doi.org/10.1037/0033-2909.103.3.324>
- Murphy, B. A., Hall, J. A., & Duong, F. (2022). It looks like construct validity, but look again: Comment on Clutterbuck et al. (2021) and recommendations for test developers in the broad “empathy” domain. *Psychological Assessment*, 34, 397–404. <https://doi.org/10.1037/pas0001063>
- Nalbant, A. (2022). Predictors of social functioning in people with antisocial personality disorder. *Dusunen Adam: The Journal of Psychiatry and Neurological Sciences*. <https://doi.org/10.14744/DAJPNS.2022.00180>
- Neumann, D. L., & Westbury, H. R. (2011). The psychophysiological measurement of empathy. In *Psychology of Empathy* (pp. 119–142). Scopus.
- Newbury-Helps, J., Feigenbaum, J., & Fonagy, P. (2017). Offenders With Antisocial Personality Disorder Display More Impairments in Mentalizing. *Journal of Personality Disorders*, 31(2), 232–255. https://doi.org/10.1521/pedi_2016_30_246

- 1 Oakley, B. F. M., Brewer, R., Bird, G., & Catmur, C. (2016). Theory of mind is not theory of
2 emotion: A cautionary note on the Reading the Mind in the Eyes Test. *Journal of*
3 *Abnormal Psychology*, 125(6), 818–823. <https://doi.org/10.1037/abn0000182>
- 4 Oswald, P., Saloppé, X., Ducro, C., Macquet, D., Cornu, P.-J., Pham, T., & Delaunoy, B. (2017).
5 Caractéristiques cliniques d'une population internée: Un cas particulier, l'établissement
6 de défense sociale « Les Marronniers » à Tournai (Belgique). *L'Encéphale*, 43(3), 229–
7 234. <https://doi.org/10.1016/j.encep.2015.10.002>
- 8 Paulhus, D. L., & Williams, K. M. (2002). The Dark Triad of personality: Narcissism,
9 Machiavellianism, and psychopathy. *Journal of Research in Personality*, 36(6), 556–563.
10 [https://doi.org/10.1016/S0092-6566\(02\)00505-6](https://doi.org/10.1016/S0092-6566(02)00505-6)
- 11 Pham, T. H., & Saloppé, X. (2010). PCL-R Psychopathy and its Relation to DSM Axis I and II
12 Disorders in a Sample of Male Forensic Patients in a Belgian Security Hospital.
13 *International Journal of Forensic Mental Health*, 9(3), 205–214.
14 <https://doi.org/10.1080/14999013.2010.517255>
- 15 Quesque, F., & Rossetti, Y. (2020). What Do Theory-of-Mind Tasks Actually Measure? Theory
16 and Practice. *Perspectives on Psychological Science*, 15(2), 384–396.
17 <https://doi.org/10.1177/1745691619896607>
- 18 Sheehan, D. V. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): The
19 Development and Validation of a Structured Diagnostic Psychiatric Interview for DSM-
20 IV and ICD-10. *J Clin Psychiatry*.
- 21 Silani, G., Lamm, C., Ruff, C. C., & Singer, T. (2013). Right Supramarginal Gyrus Is Crucial to
22 Overcome Emotional Egocentricity Bias in Social Judgments. *Journal of Neuroscience*,
23 33(39), 15466–15476. <https://doi.org/10.1523/JNEUROSCI.1488-13.2013>

- 1 Song, Z., Jones, A., Corcoran, R., Daly, N., Abu-Akel, A., & Gillespie, S. M. (2023).
2 Psychopathic traits and theory of mind task performance: A systematic review and meta-
3 analysis. *Neuroscience & Biobehavioral Reviews*, 151, 105231.
4 <https://doi.org/10.1016/j.neubiorev.2023.105231>
- 5 Stone, V. E., Baron-Cohen, S., & Knight, R. T. (1998). Frontal Lobe Contributions to Theory of
6 Mind. *Journal of Cognitive Neuroscience*, 10(5), 640–656.
7 <https://doi.org/10.1162/089892998562942>
- 8 Sulzer, S. H., Feinstein, N. W., & Wendland, C. L. (2016). Assessing empathy development in
9 medical education: A systematic review. *Medical Education*, 50(3), 300–310.
10 <https://doi.org/10.1111/medu.12806>
- 11 Surtees, A., Samson, D., & Apperly, I. (2016). Unintentional perspective-taking calculates
12 whether something is seen, but not how it is seen. *Cognition*, 148, 97–105.
13 <https://doi.org/10.1016/j.cognition.2015.12.010>
- 14 Trull, T. J., Jahng, S., Tomko, R. L., Wood, P. K., & Sher, K. J. (2010). Revised NESARC
15 Personality Disorder Diagnoses: Gender, Prevalence, and Comorbidity with Substance
16 Dependence Disorders. *Journal of Personality Disorders*, 24(4), 412–426.
17 <https://doi.org/10.1521/pedi.2010.24.4.412>
- 18 Van Dongen, J. D. M. (2020). The Empathic Brain of Psychopaths: From Social Science to
19 Neuroscience in Empathy. *Frontiers in Psychology*, 11, 695.
20 <https://doi.org/10.3389/fpsyg.2020.00695>
- 21 van Langen, M. A. M., Wissink, I. B., van Vugt, E. S., Van der Stouwe, T., & Stams, G. J. J. M.
22 (2014). The relation between empathy and offending: A meta-analysis. *Aggression and*
23 *Violent Behavior*, 19(2), 179–189. <https://doi.org/10.1016/j.avb.2014.02.003>

- 1 Vanderveren, E., Bogaerts, A., Claes, L., Luyckx, K., & Hermans, D. (2021). Narrative
2 Coherence of Turning Point Memories: Associations With Psychological Well-Being,
3 Identity Functioning, and Personality Disorder Symptoms. *Frontiers in Psychology*, 12.
4 <https://www.frontiersin.org/articles/10.3389/fpsyg.2021.623903>
- 5 Velotti, P., Garofalo, C., Dimaggio, G., & Fonagy, P. (2019). Mindfulness, Alexithymia, and
6 Empathy Moderate Relations Between Trait Aggression and Antisocial Personality
7 Disorder Traits. *Mindfulness*, 10(6), 1082–1090. [https://doi.org/10.1007/s12671-018-](https://doi.org/10.1007/s12671-018-1048-3)
8 1048-3
- 9 Vicenzutto, A., Saloppé, X., Ducro, C., Milazzo, V., Lindekens, M., & Pham, T. H. (2018).
10 Forensic Inpatients with Low IQ and Psychiatric Comorbidities: Specificity and
11 Heterogeneity of Psychiatric and Social Profiles. *International Journal of Forensic*
12 *Mental Health*, 17(3), 272–284. <https://doi.org/10.1080/14999013.2018.1504352>
- 13 Werner, K. B., Few, L. R., & Bucholz, K. K. (2015). Epidemiology, Comorbidity, and
14 Behavioral Genetics of Antisocial Personality Disorder and Psychopathy. *Psychiatric*
15 *Annals*, 45(4), 195–199. <https://doi.org/10.3928/00485713-20150401-08>
- 16 Yoder, K. J., Harenski, C. L., Kiehl, K. A., & Decety, J. (2022). Psychopathic traits modulate
17 functional connectivity during pain perception and perspective-taking in female inmates.
18 *NeuroImage: Clinical*, 34, 102984. <https://doi.org/10.1016/j.nicl.2022.102984>
- 19 x. (2024). Self-other distinction and schizotypy: Affect sharing and alexithymia in the prediction
20 of socially anxious and avoidant traits. *Personality Disorders: Theory, Research, and*
21 *Treatment*. Under review.
- 22