

Irony detection in patients with borderline personality disorder: an experimental study examining schizotypal traits, response biases and empathy

Anne Katrin Felsenheimer (✉ annefelsenheimer@googlemail.com)

Universitätsklinikum Tübingen

Carolin Kieckhäfer

LVR Klinik Düsseldorf

Alexander Michael Rapp

Universitätsklinikum Tübingen

Research Article

Keywords: schizotypal personality, sarcasm, negativity bias, social cognition, pragmatic language, mentalization, nonliteral

Posted Date: May 16th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1605408/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background: In verbal irony we often convey meanings that oppose the literal words. To look behind these words, we need to integrate perspectives of ourselves, others, and their beliefs about us. Although patients with borderline personality disorder (BPD) experience problems in social cognition and schizotypal symptoms, research on irony comprehension mainly focused on the schizophrenic spectrum. Accounting for possible negative biases in BPD, the current study examined the detection of praising and critical irony in a text message interface.

Methods: The cross-sectional study included 30 patients and 30 matched controls, who completed measures of cognitive and affective empathy (Interpersonal Reactivity Index, IRI), schizotypal (Schizotypal Personality Questionnaire; SPQ), and borderline symptoms (Borderline Symptom List; BSL-23) and the irony detection task. The irony task contained critical and praising remarks embedded in a text message interface. Asking for literality (ironic vs. literal) and intention ratings (critical to praising) of the stimuli, it allowed to analyze the sensitivity of literality detection as well as implicit and explicit response biases in a signal detection framework.

Results: Borderline symptoms explained lower sensitivity for the detection of literal and ironic statements across groups. While HC showed a negativity bias on an implicit level, BPD perceived praising remarks less praising when explicitly asked, but no difference on critical remarks. Neither empathy, nor schizotypy explained any outcomes.

Conclusions: This was the first study to show lower detection of verbal irony in patients with BPD. While patients were less biased and more neutral in implicit measurements of perceived intention, they perceived praising remarks as less positive on explicit measurements. Results stress the potential resource of non-verbal therapies to establish an appreciating therapeutical environment and contribute to difficulties in self-other distinction in BPD.

Introduction

The psychopathology of borderline personality disorder (BPD) is rooted in social interaction. In line with this, research on BPD has focused increasingly on the inferences people draw from these social interactions, so-called social cognition (1–3). One of the numerous concepts within the domain of social cognition is mentalization (4, 5). It comprises the implicit and explicit understanding of oneself and others (6) and is developed in early social interactions throughout childhood. Mentalization-based theory proposes an errant development of mentalization contributes to etiology of BPD (6, 7). In a healthy environment, the caregiver shows the child to that it is seen as an intentional being by empathically mirroring the child's expressed mental state (e.g., crying). This helps the child to internalize a coherent representation of the self and others (6). Moreover, caregivers indicate with ostensive cues (i.e., turn-taking or appropriate eye contact) that they are communicating social relevant information (8). This fosters epistemic trust, which is the general assumption that the information we receive from others is accurate and reliable. It assures us that we are not being intentionally misinformed and allows us to integrate given information in our knowledge about the world (8–10). However, individuals with BPD experience childhood adversity 13 times more often than the non-clinical adult population, especially emotional abuse and neglect (11). Such a history of maltreatment and neglect may facilitate mistrust around communication in patients with BPD, which can make it harder to believe others (10, 12). With an abusive parent, integrating given information can be dangerous (13) and those with BPD may overinterpret hostile motives when there are none. This form of 'hypermentalization' (14–22) preserves and prolongs interpersonal conflicts (12).

But to master the complexity of communication, we not only have to trust the information given, but we also have to distrust it from time to time. A prime example is verbal irony, in which vocal pitch or incongruent context information implies that the speaker intends the opposite of the literal words (23). Impairments in the comprehension of irony has been mostly demonstrated for autism (24) and schizophrenia (25–30). Being equipped with higher mistrust in the first place, it is not unlikely that patients with BPD, too, may have difficulties to decide which information to trust in irony. Above that, BPD is eponymously described as the 'border' between psychosis and neurosis (31) and individuals with BPD share cognitive biases

with schizophrenia (32), show schizotypal traits (33), and/or psychotic symptoms (34–38). Such a transdiagnostic symptomatology challenges the differential diagnostic specificity of language phenomena.

In irony, there are two causes of misinterpretation: not being able to detect the intention of the speaker, and being able to, but opting for the literal meaning regardless. Not being able to detect the intention is closely linked with mentalization. Irony requires to infer the intention from the literal words within a given context. This is why it has been mostly examined in research on social cognition (39–41) and is even used a direct measure of social cognition in video-based tasks (14, 42). The Movie for the Assessment of Social Cognition (MASC; 42) is one of them and has often been applied in BPD (18, 43–45). However, the MASC includes irony as one of several items and lacks a specific irony score. Still, Németh et al. (1) showed that in these multimodal tasks, individuals with BPD's social cognition impairments are most pronounced (1, 46, 47). Their response formats offer different interpretations of social situations (42), so selecting the right one requires the subject to explicitly compare different mental states (1). In these tasks, participants with BPD show the ability to reason about mental states, but tend to overinterpret social cues (14–22). By contrast, they show no impairments in nonverbal paradigms such as the Reading the Mind in the Eyes Task (48). This test only requires detecting an emotion based on pictures of the eye region, but not to reason about it (1). The authors concluded that the implicit detection of emotions seems to be preserved in BPD. Instead, difficulties arise when multiple perspectives need to be explicitly compared. Multiple perspectives, however, are an inevitable part of irony (39, 49, 50). And detecting irony requires a flexible shift between them - shifts that have just recently been shown to be harder for individuals with BPD (51–53).

Yet, even the full capacity to compare mental states does not necessarily guarantee that a statement will be perceived as ironic. Irony explicitly leaves the intention of the speaker ambiguous and along that room for interpretation. Individuals who tend to perceive others as dishonest may decide to ignore irony, irrespective of their ability to infer mental states. Addressing this distinction in schizophrenia, Parola et al. (27) analyzed both sensitivity (the detection of a communicative intention) and response bias (the tendency to favor a specific response) during indirect speech comprehension. Individuals with schizophrenia had equal difficulty detecting ironic, deceitful, and sincere phrasings, but tended to perceive ironic utterances more deceitful than healthy controls. Negative attribution styles are common in BPD as well (2); many individuals with BPD tend to interpret others' behavior as aggressive and hostile (54) and neutral faces as less trustworthy (55). This leaves patients in a vicious cycle of re-experiencing traumatic relationships (13, 56). Therapists are often encouraged to use a clear, unambiguous communication to avoid unintentionally reinforcing their patient's perceived threat (57). Especially, as negative biases in BPD tend to evolve specifically with ambiguous stimuli (3, 47, 58–60).

Using irony as a prime example of ambiguous language allows both pragmatic inference and attributional bias to be examined within one linguistic phenomenon. Most studies on irony comprehension, however, focus solely on ironic criticism or sarcasm (42), thereby confounding irony with an a priori negative bias. Analyzing both praising and critical irony bypasses positive testing and allows interpretation errors to be analyzed without overtly asking for them. For example, Kieckhäfer et al. (61) examined how the detection of ironic and literal praising and critical relate to borderline and schizotypal traits in healthy adults. In their study, both traits were associated with lower detection accuracy, though each set of traits had differing error patterns. In line with Parola et al.'s findings in schizophrenia (27), individuals with higher schizotypy interpreted the stimuli more mocking: They indicated literal praise as ironic critique and ironic praise as literal critique. In contrast, individuals with high borderline traits only made errors identifying ironic remarks and this was regardless of the intention.

We applied Kieckhaefer et al.'s (61) paradigm, for the first time, on participants diagnosed with BPD. We compared the detection of literality (ironic vs. literal) and implicit response biases within a signal detection theory (SDT) framework, as well as explicit ratings of the perceived intention (critical to praising) with healthy controls (HC). In accordance with findings on healthy adults with borderline symptoms (61), we hypothesized that participants with BPD would have more difficulty differentiating ironic and literal utterances. We further assumed that negativity biases would emerge in a clinically relevant BPD symptomatology. Thus, in contrast to Kieckhäfer et al.'s results (61), we expected participants with BPD to interpret ironic praise and literal criticism literally, and ironic criticism and literal praise ironically. In line with this, we predicted BPD

participants would rate critical remarks as more critical and praising remarks as less praising. Lastly, we explored the relationship between BPD symptoms, schizotypal traits, and mentalization, assessed via affective and cognitive empathy.

Methods

30 participants with BPD were recruited at the University Hospital of Tuebingen, Germany. General exclusion criteria were acute or anamnestic substance abuse or dependence, bipolar disorder, psychotic disorders, severe episodes of major depression, and neurological diseases. Inclusion criteria involved normal or corrected-to-normal vision, age between 18–55, native German speaking status, and a clinical diagnosis of BPD for the patient group. A trained clinician assessed the DSM-IV criteria according to Structured Clinical Interview for DSM IV II(62) and comorbidities according to SCID I (63). The study protocol was approved by the ethics committee of the Medical Faculty of the University of Tuebingen and carried out according to the Declaration of Helsinki. All participants provided written informed consent and received monetary compensation.

A group of 30 healthy controls (HC), was matched for age, verbal intelligence according to the multiple-choice vocabulary test (MWT; 64), gender, and educational level. Both groups filled out the short version of the Borderline Symptom List (BSL-23; 65) and the German version of the Schizotypal Personality Questionnaire (SPQ; 66,67). For the evaluation of cognitive and affective empathy, the Interpersonal Reactivity Index (IRI; 68) was used as a German short version (69). The IRI is a self-report instrument comprising two cognitive subscales (perspective taking, fantasy) and two affective subscales (empathic concern, personal distress).

After consenting to participate, demographics were assessed in paper-pencil format. Then, participants completed the irony paradigm and self-report instruments on a computer in a quiet, distraction-free room. The stimuli were the same as in Kieckhäfer et al. (61); test construction and development are explained in detail there. Each trial consisted of a videotaped context story introducing a character in a café and subsequent message exchanges. According to the narrative, participants saw text messages containing a context sentence and a reaction to that message by the protagonist of the video (see additional file 1). The message was either ironic praise (IP), ironic criticism (IC), literal praise (LP), or literal criticism (LC). In ironic stimuli, the intended meaning opposed the literal meaning. For example, IP had a praising intention by way of critique (“I went running today” “You are so lazy.”). Videos varied in the degree of proximal perspective, and were either addressed directly by the protagonist (2nd person) or observed by the protagonist talking to a neutral other (3rd person). The protagonist’s answers were to be scored on their literality (ironic vs. literal) in a binary response format, and their perceived intention (criticism vs. praise) on a five-point Likert scale (see additional file 2). Each trial comprised five items per condition (20 items total). Participants completed two test versions, with one perspective each. Summation of correct identified items lead to a total maximum score of 10 correct responses per condition (IP, IC, LP, LC) for both test versions.

We applied SDT to quantify sensitivity (d') and response biases (β). As in SDT designs, the irony task required a binary label of literality (literal vs. ironic), which could be compared to the presence or absence of a signal (irony present vs. irony absent), resulting in four logical outcomes (Table 1): hit (choosing ironic in an ironic stimulus), false alarm (choosing ironic in a literal stimulus), miss (choosing literal in an ironic stimulus), and correct rejection (choosing literal in a literal stimulus). Each category was assigned a likelihood ratio. For instance, the hit rate represents the proportion of ironic stimuli to which the participant responded “ironic,” and false alarm rate denotes the proportion of literal trials to which the participant responded “ironic”. Unlike the mere number of correct responses, SDT’s measure of sensitivity reflects the probability of identifying the intention of the stimulus while avoiding false alarms, and corresponds to the Z-value of the hit rate minus the false-alarm rate.

Table 1
Signal detection theory matrix with possible outcomes for each contrastive pair of stimuli.

		response	
		ironic	literal
stimulus pair			
IC vs. LP	IC (irony present)	hit	miss
	LP (irony absent)	false alarm	correct rejection
IP vs. LC	IP (irony present)	hit	miss
	LC (irony absent)	false alarm	correct rejection
<i>Notes.</i> IC = ironic criticism; IP = ironic praise; LC = literal criticism; LP = literal praise.			

SDT further accounts for the response bias β : a systematic criterion when a signal is considered as present. It can capture the tendency of an individual to interpret statements either as ironic or literal. An individual who tends to interpret statements as “ironic” shows high hits for ironic (IC and IP), but high false alarms in literal stimuli (LC and LP). An unbiased observer’s β is close to 1. With a tendency to respond “ironic” (liberal criterion), β approaches 0. Vice versa - with the tendency to choose “literal” (conservative criterion), β exceeds 1. d' and β were computed with the R package Psycho. The binary answer format (ironic vs. literal) and definition of irony as the opposite of literal meaning resulted in two corresponding conditions (IP vs. LC; IC vs. LP). Specifically, in an IC stimulus (“I am too late.” “You are so reliable.”), the detection of the correct literality (i.e., “ironic”) requires detecting the critical intention, despite the literal praise. The same holds true for IP and LC. For each participant, we calculated the hits, false alarms, misses, and correct rejections for both matching pairs.

Then, we applied linear mixed effect models in R with the lme4 package using d' , β , and ratings of perceived intention as respective outcome; group (HC vs. BPD) and intention (praise vs. criticism) as sum-coded fixed effects; age and verbal intelligence as continuous covariates; gender as a categorical covariate; and random effects by participant. Post-hoc tests with adjusted p-values were carried out with Tukey’s test. Based on the stimulus design, misclassifying literality causes perception of the opposite intention (e.g., ironic praise as literal criticism). Thus, perceived intention was estimated by the mean rating of items correctly identified as ironic or literal.

For each model, the impact of borderline symptoms, schizotypal symptoms, and empathy scales was analyzed. Model fits were estimated hierarchically, starting out with the null model, then adding borderline and schizotypal symptoms, and finally IRI subscales, as fixed effects. Models were compared via Likelihood-ratio tests using the anova function.

Similar to previous results (61), the perspective of the speaker had no effect on detection performance in a preceding repeated measure Analysis of Variance (rmANOVA, see supplementary table 1). Thus, the conditions were not included in analysis.

Results

Groups did not differ significantly in age ($t(58) = -.812, p = .420$), gender ($Z = .417, p = .519$), educational level ($Z = -1.736, p = .083$), or verbal intelligence ($t(58) = -1.062, p = .293$). Patients with BPD had significantly more borderline symptoms ($t(38.46) = -8.971, p < .001$) and personal distress ($t(58) = -6.215, p < .001$), as well as lower perspective taking ($t(58) = 2.871, p = .006$) and more schizotypal symptoms ($t(58) = -8.662, p < .001$). A detailed sample description can be found in Table 2.

Table 2
Means (M) and standard deviations (SD) of demographic and psychometric data.

	BPD (n = 30)		HC (n = 30)		p
	M	SD	M	SD	
demographics					
age (years)	29.27	9.03	27.20	10.03	.420 ^a
gender (female/male)	25/5		23/7		.519 ^b
education (median/IQR)	4.00	1.25	4.00	0.00	.083 ^c
verbal intelligence	28.51	4.07	28.89	3.71	.293 ^a
questionnaires ^a					
BSL-23	2.28	1.02	0.39	0.40	< .001
IRI					
personal distress	15.87	3.01	10.60	3.53	< .001
empathetic concern	15.63	2.47	14.73	2.60	.174
perspective taking	13.30	2.74	15.27	2.56	.006
fantasy	13.80	3.94	14.47	3.14	.472
SPQ	34.7	13.5	15.5	10.6	< .001
perceived intention					
IC	2.25	0.62	2.17	0.46	
IP	3.56	0.69	4.01	0.47	
LC	1.94	0.49	1.81	0.42	
LP	4.52	0.37	4.70	0.25	
sensitivity (d')					
IC vs. LP	1.11	0.31	1.17	0.21	
IP vs. LC	0.89	0.31	1.07	0.22	
response bias (β)					
IC vs. LP	1.01	0.09	0.95	0.09	
IP vs. LC	1.07	0.11	1.13	0.09	
<p><i>Notes.</i> ^a independent sample t-test, ^b Pearson-Chi-Quadrat, ^c Mann-Whitney-U-Test, ^d Welch-Test. HC = healthy controls; BPD = borderline personality disorder; BSL-23 = Borderline Symptom List; IRI = interpersonal reactivity index; SPQ = schizotypal personality questionnaire; IC = ironic criticism; IP = ironic praise; LC = literal criticism; LP = literal praise.</p>					

Final models are depicted in Table 3. There was no effect of gender, age, or verbal IQ. Despite possible ceiling effects, patients with BPD ($M = 1.00$, $SD = 0.32$) showed significantly less sensitivity than HC ($M = 1.12$, $SD = 0.22$) in differentiating ironic and literal statements ($t(58) = 2.184$, $p = .033$), regardless of the intention (see Fig. 1). Above groups, sensitivity was higher for IC vs. LP ($M = 1.14$, $SD = .27$) than IP vs. LC ($M = .98$, $SD = .28$; $t(58) = 4.22$, $p < .001$).

Table 3

Analysis of deviance table (Type II Wald chi-square tests) for the linear mixed effect models with sum-coded contrasts and random intercepts by subject of sensitivity (left) and response bias (right) including borderline symptoms.

	sensitivity				response bias			
	$d' \sim \text{group} * \text{intention} + \text{bsl} + \text{age} + \text{gender} + \text{iq} + (1 \text{ID})$				$\beta \sim \text{group} * \text{intention} + \text{age} + \text{gender} + \text{iq} + (1 \text{ID})$			
fixed effects	<i>b</i>	χ^2	<i>df</i>	<i>p</i>	<i>b</i>	χ^2	<i>df</i>	<i>p</i>
group	0.00	0.01	1	.752	0.00	0.01	1	.923
intention	0.08	17.81	1	< .001***	-0.06	51.86	1	< .001***
group*intention	-0.03	2.31	1	.129	-0.03	10.90	1	< .001***
age	0.00	0.24	1	.623	0.11	0.73	1	.736
gender	0.05	0.62	1	.432	0.04	4.11	1	.042*
IQ	0.01	1.53	1	.216	-0.00	0.58	1	.446
BSL	-0.08	4.12	1	.042*				

Notes. BSL = score on borderline symptom list 23.

For response bias, there was a significant interaction of group on intention. However, post-hoc pairwise comparisons of group by level of intention indicated group differences in β for both IC vs. LP ($t(116) = -2.313, p = .023$) and IP vs. LC ($t(116) = 2.321, p = .022$). On a descriptive level, β tended to be closer to 1 in BPD (see Table 1), indicating that BPD participants were almost unbiased. In contrast, HC showed a lower β in ironic criticism, corresponding with a tendency to interpret an answer as ironic in IC and LP and implying a tendency to interpret stimuli as mocking. The same negativity bias was evident in the other pair, with HC having a higher β in IP vs. LC and a tendency to choose literal. There was a significant effect of gender, with males having higher d' than females ($t(55) = -2.027, p = .048$).

Borderline symptoms significantly improved model fit for d' ($\chi^2(1) = 5.497, p = .019$), with a significant effect on d' diminishing the effect of group. Neither SPQ ($\chi^2(1) = .187, p = .633$), nor IRI subscales ($\chi^2(4) = 2.302, p = .680$) improved model fit. For β , neither BSL ($\chi^2(1) = 0.011, p = .917$), nor SPQ ($\chi^2(1) = 0.517, p = .472$) or IRI scales ($\chi^2(1) = 0.064, p = .999$) improved model fit.

Figure 2. Response bias for the comparisons of ironic criticism (IC) with literal praise (LP) and ironic praise (IP) with literal criticism (LC). An unbiased β corresponds to 1, an β approaching zero a tendency to choose ironic, a β increasing over one a tendency to choose literal as response.

In the last step we analyzed the ratings of perceived intention (Table 4). There was a significant interaction of group*intention, with clinical participants perceiving praising remarks as less praising ($t(172) = 3.480, p < .001$), but no difference in perception of critical remarks ($t(172) = -1.133, p = .259$). Post-hoc comparisons for interaction of intention*literal indicated all pairwise comparisons to be significant (all $p < .0001$), confirming previous findings that ironic remarks were perceived as less praising (IP: $M = 2.21, SD = .63$ vs. LP: $M = 4.61, SD = .33$) and less critical (IC: $M = 2.21, SD = .54$ vs. LC: $M = 1.87, SD = .46$) in both groups. Again, neither BSL ($\chi^2(1) = 0.472, p = .491$), SPQ ($\chi^2(1) = .740, p = .387$), nor IRI scales ($\chi^2(1) = 5.017, p = .028$) improved model fit.

Table 4
 Analysis of deviance table (Type II Wald chi-square tests) for the linear mixed model of perceived intention with sum-coded contrasts and random intercepts by subject.

perceived intention				
rating ~ group*intention*literality + (1 ID)				
fixed effects	<i>estimate</i>	<i>X</i> ²	<i>df</i>	<i>p</i>
between-subject				
group	0.053	2.72	1	.098
group*literality	0.042	1.74	1	0.187
group*intention	-0.103	10.77	1	< .001***
group*literality*intention	-0.027	0.75	1	.386
within-subject				
literality	-0.123	15.15	1	< .001***
intention	-1.076	1158.61	1	< .001***
intention*literality	0.290	84.45	1	< .001***

Discussion

This was the first study to examine irony comprehension among individuals with BPD. Participants were presented with both ironic and literal text messages varying in praising and critical intention. Within a signal detection framework, we assessed negativity biases and the ability to discriminate literal from ironic remarks. Biases were distinguished on two levels: implicit tendencies measured in the choice of the literalness of the statement (ironic vs. literal) and explicit ratings of perceived intention (critical to praising).

Participants with BPD exhibited more difficulty differentiating literal from ironic remarks than HC. Yet, group differences did not vary with critical or praising content, showing that it was the literality of the stimulus, not the intention, affecting performance. For both groups, ironic praise was harder to detect than ironic criticism, replicating that ironic criticism is easier to process (70, 71), mostly because it is much more common (71–73). A reduced sensibility was explained by borderline pathology beyond groups, confirming findings among healthy adults with borderline traits in a clinical sample (61) and corroborating the idea of dimensional approaches to personality disorders (74). The current results are commensurate with those in social cognition paradigms using sarcasm as a stimulus (15, 22, 44). For the first time, these impairments have been confirmed with respect to verbal irony.

Other forms of nonliteral language, such as metaphors, have recently been shown to be preserved in BPD (75). This is of particular importance, as metaphor comprehension is commonly impaired in schizophrenia (76) with whom BPD patients share symptoms (32, 33). In addition to emphasizing that linguistic phenomena are subject to different cognitive structures, this shows that psychopathologies may be subject to them as well. Both metaphor and irony require an abstraction from the literal words, but irony further demands to integrate multiple mental states (8, 77, 78). Ironists do not intend to deceive, but seek duplicitous understanding. As such, irony proves particularly challenging for epistemic trust, as the listener must decide which of the conflicting pieces of information to trust and which to question. In our study, errors in the detection of irony implied that in some instances patients decided to stick to the literal meaning, even when an incongruence between context and target sentence suggested otherwise. A lack of epistemic trust may make it more difficult for individuals with BPD to decide which information is relevant in an ambiguous context (10). As a result, they may adhere to one perspective (16, 79) and choose a context-inappropriate interpretation. Indeed, shifts in the representation of the self and of others have long

been deemed problematic in BPD (31). Empirically, patients experience difficulties alternating between egocentric and altercentric perspectives with face-morphing tasks (52) and show overlapping self-other boundaries on a bodily and cognitive level (51, 52). Accordingly, in our study, patient's personal distress in response to others' emotions was higher and cognitive perspective-taking lower than those of HC, replicating previous findings on self-reported empathy scales in BPD (80–82). However, both were unrelated to outcomes in our study of irony. Future studies should include more complex social cognition paradigms that may be more commensurate with metacognitive processes than with self-ratings (46), and should compare speech varying in self-other representation (e.g., deceit and faux pas).

Contrary to our expectations, HC (and not BPD) tended to interpret stimuli critically when looking at implicit response biases. These were measured in participants' interpretation of a remark as literal or ironic. Interpreting literal praise ironically HC ascribed negative intent to literal praise ("I have an A in my test" "You are clearly not smart"), while ironic criticism was interpreted as literal criticism. The same negativity bias was evident in interpretations of ironic praise and literal criticism: HC tended to interpret these statements literally, considering ironic praise as literal critique and literal critique as such. Although individuals with BPD's difficulty to determine the literality of a statement may be explained by overattribution (83), our results did not indicate that these attributions are hostile when measured implicitly.

Response biases were further analyzed with explicit ratings of perceived intention. Both groups tended to rate ironic utterances as less praising and critical than literal ones, confirming the well-known perception of irony as "tinged" with the literal meaning (84, 85). In contrast to implicit biases, patients with BPD showed a negative interpretation bias in explicit ratings. They perceived praising remarks as less praising than HC. So far, only a limited number of studies have investigated the effect of positive social stimuli in BPD (47). Our findings are in line with BPD participants' fear of positive appraisal (86), negative ratings of appreciating video-clips (87) or self-referential information (88) and approach-avoidance behavior (89, 90). Muting the positive experience of praise has major implications, since positive feedback is a crucial part of the therapeutic process (91) and of positive interactions with others. Yet, contrary to other studies (87) and patients with BPD's heightened rejection sensitivity (92) clinical participants did not differ to HC in perception of critical remarks.

Several factors may contribute to the discrepancy between explicit and implicit biases in BPD. First, the current stimuli were written text messages. Negativity biases in BPD have mostly been found in facial emotion recognition (3, 47, 93, 94), especially when combined with other modalities (95, 96). Further, most biases accompany anger and disgust (3, 93) or neutral stimuli (3, 59, 97). Critique may be less arousing than expressions of anger. And irony is almost impossible to be neutral, as its principal function is to tacitly convey an opinion of the ironized content (50). Yet, unlike most emotion recognition paradigms (47), the current one allowed for an implicit assessment of response tendencies. In this context, our findings correspond with Kobeleva et al.'s (98) comparing approach-avoidance reactions with photographs of positive and negative expressions. In their study, patients tended to rate positive faces as less approachable, but there were no differences between BPD and HC in avoidance and approach behaviors. Similarly, Franzen et al. (99) measured participant investments in a virtual trust game without asking for evaluations of the trustees. HC considered the trustees' facial expressions and adjusted their investments accordingly, while BPD relied exclusively on the fairness of the trustees' decisions. Our results align with these findings and indicate that negative appraisals in BPD tend to emerge in contexts in which emotional ratings are salient.

We are aware of several limitations. First, our patient sample had a high verbal IQ and educational background, which may be less prevalent in BPD among the general population (100, 101). Further, patients were recruited in a specialized ward for dialectal behavioral therapy (102), which trains the differentiation between self and other and emotion regulation. This might have minimized group differences. Second, our stimuli did not contain prosody or facial expressions. However, the study focused on text messages - a conversational context that is a major part of current communication. Third, the concept of irony transcends verbal irony, such as situational irony, hyperbole or understatement (103). Fourth, we did not account for experienced abuse or neglect, which is considered to be the origin of epistemic distrust (10). Lastly, this study did not include clinical controls, leaving the question of clinical specificity to be explored. As both personality dimensions often overlap (33), it is important to analyze distinct and/or shared mechanisms of disorders in the schizophrenic spectrum and BPD. Contrary

to other studies (28, 29, 61, 104, 105), schizotypal symptoms did not explain irony detection beyond borderline symptoms in our study, although patients scored high on both. This further stresses the need to control for borderline symptoms when assessing irony comprehension in different pathologies.

Conclusions

This was the first study to provide evidence for an impaired irony detection in patients with BPD. Borderline symptoms explained this effect, but neither schizotypal traits nor empathy scales were related to outcomes of irony. While the use of ambiguous language is claimed to be restricted in therapeutic contexts with BPD patients (7, 91), the current study shows that this claim cannot be generalized to all forms of nonliteral language. With a preserved metaphor (75), but impaired irony comprehension in BPD, it seems that it is not the ambiguity of being nonliteral, but the ambiguity of the intention that imposes an obstacle for BPD. Just as irony forms a sense of collusion for those who understand (106), a patient's misunderstanding may leave them with feelings of exclusion. In BPD this may even lead to a rupture in the therapeutic relation. Patients with BPD only showed a negative bias when explicitly asked to rate perceived intention. To them, praising remarks were considered less praising. Therapists and research alike naturally focus on negative social perceptions in BPD, but our results highlight the importance of targeting the diminished beneficial effect of positive feedback as well. Both findings emphasize the relevance of a shared and open discussion of the possible inferences BPD patients may draw from social interactions. But they also stress the vital role of non-verbal, embodied treatments beyond talking therapies (12, 107). Explicit mentalization is not a fully abstract process, but inherently interwoven with implicit, bodily intersubjectivity (108, 109). Using lack of implicit biases in BPD as a resource, non-verbal therapies can provide a realm of unequivocal appreciation. They may serve to build a solid ground for the exploration of self and others in language.

Abbreviations

rmANOVA: repeated measurements analysis of variance

BPD: Borderline personality disorder

BSL-23: Borderline Symptom List – short version

DSM-IV: Diagnostic and Statistical Manual of Mental Disorders

HC: Healthy control participants

IC: Ironic criticism

IP: Ironic praise

IRI: Interpersonal reactivity index

LC: Literal praise

LP: Literal criticism

MASC: Movie for the Assessment of Social Cognition

MWT: German multiple-choice vocabulary test

SCID-II: Structured Clinical Interview for axis II personality disorders

SCID-I: Structured Clinical Interview for axis I

SDT: Signal detection theory

Declarations

Availability of data and materials

The datasets and material in the current study are available from the corresponding author upon request.

Acknowledgments

We would like to thank all the participants for their involvement in the study and both actors who participated in the production of the stimuli.

Funding

No funding was provided for the collection of this data or writing of this manuscript.

Affiliations

Department of Psychiatry and Psychotherapy, University of Tuebingen, Tuebingen, Germany

Anne K. Felsenheimer, Alexander M. Rapp

LVR hospital Düsseldorf, University of Düsseldorf, Düsseldorf, Germany

Carolin Kieckhäfer

Contributions

AF, CK and AR designed the study. AF and CF acquired the data. CF and AF prepared the data for analyzing. AF analyzed the data. AF wrote the article, which all authors reviewed and approved for publication.

Corresponding author

Correspondence to Anne K. Felsenheimer (annefelsenheimer@googlemail.com)

Ethics approval and consent to participate

The study was approved by the Ethics Commission of the Medical Faculty of the University Clinic of Tuebingen and complied with the Declaration of Helsinki. All participants provided written informed consent to participate in the study.

Consent for publication

Both actors provided written informed consent for publication of video and picture material of the irony paradigm.

Competing interests

The authors declare that they have no competing interests.

References

1. Németh N, Mátrai P, Hegyi P, Czéh B, Czopf L, Hussain A, et al. Theory of mind disturbances in borderline personality disorder: A meta-analysis. *Psychiatry Research*. 2018;270:143–53. Available from: <https://doi.org/10.1016/j.psychres.2018.08.049>

2. Roepke S, Vater A, Preißler S, Heekeren HR, Dziobek I. Social cognition in borderline personality disorder. *Frontiers in Neuroscience*. 2012;4(JAN):182. Available from: <http://journal.frontiersin.org/article/10.3389/fnbeh.2010.00182>
3. Daros AR, Zakzanis KK, Ruocco AC. Facial emotion recognition in borderline personality disorder. *Psychological Medicine*. 2012/11/13. 2013;43(9):1953–63. Available from: <https://www.cambridge.org/core/article/facial-emotion-recognition-in-borderline-personality-disorder/395D75347AA4822298CE80AEC43BFE75>
4. Brüne M. “Theory of Mind” in Schizophrenia: A Review of the Literature. *Schizophrenia Bulletin*. 2005 Jan 1;31(1):21–42. Available from: <https://doi.org/10.1093/schbul/sbi002>
5. Dimopoulou T, Tarazi FI, Tsapakis EM. Clinical and therapeutic role of mentalization in schizophrenia - A review. *CNS Spectrums*. 2017;22(6):450–62. Available from: <https://doi.org/10.1017/S1092852916000687>
6. Fonagy P, Bateman A. The development of borderline personality disorder - A mentalizing model. *Journal of Personality Disorders*. 2008;22(1):4–21. Available from: <https://doi.org/10.1521/pedi.2008.22.1.4>
7. Fonagy P, Bateman AW. Attachment Theory and Mentalization-Oriented Model of Borderline Personality Disorder. In: *The American Psychiatric Publishing textbook of personality disorders*. Arlington, VA, US: American Psychiatric Publishing, Inc; 2005. p. 187–207. Available from: <https://ezproxy.lib.uwm.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2005-05013-012&site=ehost-live>
8. Sperber D, Wilson D. Pragmatics, Modularity and Mind-reading. *Mind & Language*. 2002;17(1-2):3–23. Available from: <https://doi.org/10.1111/1468-0017.00186> TS - CrossRef
9. Sperber D, Clément F, Heintz C, Mascaro O, Mercier H, Origg G, et al. Epistemic vigilance. *Mind and Language*. 2010;25(4):359–93.
10. Fonagy P, Luyten P, Allison E. Epistemic petrification and the restoration of epistemic trust: A new conceptualization of borderline personality disorder and its psychosocial treatment. *Journal of Personality Disorders*. 2015;29(5):575–609. Available from: <https://doi.org/10.1521/pedi.2015.29.5.575>
11. Porter C, Palmier-Claus J, Branitsky A, Mansell W, Warwick H, Varese F. Childhood adversity and borderline personality disorder: a meta-analysis. *Acta Psychiatrica Scandinavica*. 2020 Jan 1 [cited 2021 Sep 1];141(1):6–20. Available from: <https://onlinelibrary.wiley.com/doi/full/10.1111/acps.13118>
12. Fonagy P, Luyten P, Allison E, Campbell C. What we have changed our minds about: Part 2. Borderline personality disorder, epistemic trust and the developmental significance of social communication. *Borderline Personality Disorder and Emotion Dysregulation*. 2017 Apr 11 [cited 2022 Feb 21];4(1):1–12. Available from: <https://bpded.biomedcentral.com/articles/10.1186/s40479-017-0062-8>
13. Fonagy P. Attachment and Borderline Personality Disorder. *J Am Psychoanal Assoc*. 2000 Aug 1;48(4):1129–46. Available from: <https://doi.org/10.1177/00030651000480040701>
14. Sharp C, Ha C, Carbone C, Kim S, Perry K, Williams L, et al. Hypermentalizing in adolescent inpatients: Treatment effects and association with borderline traits. *Journal of Personality Disorders*. 2013;27(SPL.ISS.1):3–18. Available from: http://sfx.ucl.ac.uk/sfx_local?sid=OVID:psycdb&id=pmid:&id=doi:10.1521%2Fpedi.2013.27.1.3&issn=0885-579X&isbn=&volume=27&issue=1&page=3&pages=3-18&date=2013&title=Journal+of+Personality+Disorders&atitle=Hypermentalizing+in+adolescent+inpatients%3A+Treath
15. Sharp C, Pane H, Ha C, Venta A, Patel AB, Sturek J, et al. Theory of mind and emotion regulation difficulties in adolescents with borderline traits. *J Am Acad Child Adolesc Psychiatry*. 2011;50(6):563–573.e1. Available from: <https://doi.org/10.1016/j.jaac.2011.01.017>
16. Sharp C, Vanwoerden S. Hypermentalizing in borderline personality disorder: A model and data. *Journal of Infant, Child, and Adolescent Psychotherapy*. 2015;14(1):33–45. Available from: <https://doi.org/10.1080/15289168.2015.1004890> TS - CrossRef
17. Penner F, McLaren V, Leavitt J, Akca OF, Sharp C. Implicit and Explicit Mentalizing Deficits in Adolescent Inpatients: Specificity and Incremental Value of Borderline Pathology. https://doi.org/10.1521/pedi_2019_33_463. 2020 Oct 12 [cited 2021 Aug 22];34:64–83. Available from: https://guilfordjournals.com/doi/abs/10.1521/pedi_2019_33_463

18. Vaskinn A, Antonsen BT, Fretland RA, Dziobek I, Sundet K, Wilberg T. Theory of mind in women with borderline personality disorder or schizophrenia: differences in overall ability and error patterns. *Frontiers in Psychology*. 2015;6:1239. Available from: <https://doi.org/10.3389/fpsyg.2015.01239>
19. Andreou C, Kelm L, Bierbrodt J, Braun V, Lipp M, Yassari AH, et al. Factors contributing to social cognition impairment in borderline personality disorder and schizophrenia. *Psychiatry Research*. 2015;229(3):872–9. Available from: <https://doi.org/10.1016/j.psychres.2015.07.057>
20. Ortega-Díaz E, García-Campos J, Rico-Gomis JM, Cuesta-Moreno C, Palazón-Bru A, Estañ-Cerezo G, et al. Social cognition and social functioning in people with borderline personality disorder and their first-degree relatives. *PeerJ*. 2020 Oct 30 [cited 2021 Sep 6];8:e10212. Available from: <https://peerj.com/articles/10212>
21. Normann-Eide E, Antonsen BT, Kvarstein EH, Pedersen G, Vaskinn A, Wilberg T. Are Impairments in Theory of Mind Specific to Borderline Personality Disorder? https://doi.org/101521/pedi_2019_33_417. 2020 Dec 2 [cited 2021 Sep 6];34(6):827–41. Available from: https://guilfordjournals.com/doi/abs/10.1521/pedi_2019_33_417
22. Somma A, Ferrara M, Terrinoni A, Frau C, Ardizzone I, Sharp C, et al. Hypermentalizing as a marker of borderline personality disorder in Italian adolescents: a cross-cultural replication of Sharp and colleagues' (2011) findings. *Borderline Personality Disorder and Emotion Dysregulation* 2019 6:1. 2019 Apr 10 [cited 2021 Aug 22];6(1):1–4. Available from: <https://bpded.biomedcentral.com/articles/10.1186/s40479-019-0104-5>
23. Lausberg H. *Handbuch der literarischen Rhetorik*. 3. Aufl. Stuttgart: Steiner; 1960. S. [605]-957. Available from: <http://www.ub.univie.ac.at/katscans/NKN009%5C0043%5C00000098.gif%5Cnhttp://ubdata.univie.ac.at/AC01079376>
24. Kalandadze T, Norbury C, Nærland T, Næss KAB. Figurative language comprehension in individuals with autism spectrum disorder: A meta-analytic review. *Autism*. 2018;22(2):99–117. Available from: <https://doi.org/10.1177/1362361316668652>
25. Langdon R, Davies M, Coltheart M. Understanding Minds and Understanding Communicated Meanings in Schizophrenia. *Mind & Language*. 2002;17(1-2):68–104. Available from: <https://doi.org/10.1111/1468-0017.00190> TS - CrossRef
26. Mo S, Su Y, Chan RCK, Liu J. Comprehension of metaphor and irony in schizophrenia during remission: The role of theory of mind and IQ. *Psychiatry Research*. 2008;157(1–3):21–9.
27. Parola A, Brasso C, Morese R, Rocca P, Bosco FM. Understanding communicative intentions in schizophrenia using an error analysis approach. *npj Schizophrenia* 2021 7:1. 2021 Feb 26 [cited 2021 Aug 11];7(1):1–9. Available from: <https://www.nature.com/articles/s41537-021-00142-7>
28. Rapp AM, Mutschler DE, Wild B, Erb M, Lengsfeld I, Saur R, et al. Neural correlates of irony comprehension: The role of schizotypal personality traits. *Brain and Language*. 2010;
29. Rapp AM, Langohr K, Mutschler DE, Klingberg S, Wild B, Erb M. Isn't it ironic? Neural Correlates of Irony Comprehension in Schizophrenia. *PLoS ONE*. 2013;
30. Varga E, Simon M, Tényi T, Schnell Z, Hajnal A, Orsi G, et al. Irony comprehension and context processing in schizophrenia during remission - A functional MRI study. *Brain and Language*. 2013;
31. Kernberg OF. *Borderline-Störungen und pathologischer Narzissmus*. 4. Aufl.,. Literatur der Psychoanalyse. Frankfurt am Main: Suhrkamp; 1980. 438 TS-GBV Gemeinsamer Bibliotheksverbund M4-. Available from: <http://gso.gbv.de/DB=2.1/PPNSET?PPN=040669343>
32. Puri P, Kumar D, Muralidharan K, Kishore MT. Individuals with Borderline Personality Disorder manifest cognitive biases implicated in psychosis. *Psychiatry Research*. 2018;267:414–9. Available from: <https://doi.org/10.1016/j.psychres.2018.06.040>
33. Kwapil TR, Clark HE, Rbeiz KS, Bathery AJ, Kemp KC, Barrantes-Vidal N. Association of positive, negative, and disorganized schizotypy with cluster a, borderline, and avoidant personality disorders and traits. *Personality Disorders: Theory, Research, and Treatment*. 2021;

34. Cavelti M, Thompson K, Chanen AM, Kaess M. Psychotic symptoms in borderline personality disorder: developmental aspects. *Current Opinion in Psychology*. 2021 Feb 1;37:26–31.
35. D'Agostino A, Rossi Monti M, Starcevic V. Psychotic symptoms in borderline personality disorder: an update. *Curr Opin Psychiatry*. 2019;32(1):22–6. Available from: <https://doi.org/10.1097/YCO.0000000000000462>
36. Moritz S, Schilling L, Wingenfeld K, Köther U, Wittkind C, Terfehr K, et al. Psychotic-like cognitive biases in borderline personality disorder. *Journal of Behavior Therapy and Experimental Psychiatry*. 2011 Sep 1;42(3):349–54.
37. Slotema CW, Daalman K, Blom JD, Diederer KM, Hoek HW, Sommer IEC. Auditory verbal hallucinations in patients with borderline personality disorder are similar to those in schizophrenia. *Psychological Medicine*. 2012/02/16. 2012;42(9):1873–8. Available from: <https://www.cambridge.org/core/article/auditory-verbal-hallucinations-in-patients-with-borderline-personality-disorder-are-similar-to-those-in-schizophrenia/40E1B1B007825F7EFE3B3F750A133AE2>
38. Slotema CW, Blom JD, Deen M, Niemantsverdriet MBA, van der Gaag M, Hoek HW, et al. Negative Beliefs about Voices in Patients with Borderline Personality Disorder Are Associated with Distress: A Plea for Cognitive-Behavioural Therapy? *Psychopathology*. 2017;50(4):255–61. Available from: <https://www.karger.com/DOI/10.1159/000477669>
39. Winner E, Brownell H, Happé F, Blum A, Pincus D. Distinguishing lies from jokes: Theory of mind deficits and discourse interpretation in right hemisphere brain-damaged patients. *Brain and Language*. 1998;
40. Happé FGE. An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental Disorders*. 1994;24(2):129–54. Available from: <https://doi.org/10.1007/BF02172093>
41. Bosco FM, Gabbatore I. Sincere, deceitful, and ironic communicative acts and the role of the theory of mind in childhood. *Frontiers in Psychology*. 2017;8:21. Available from: <https://doi.org/10.3389/fpsyg.2017.00021>
42. Dziobek I, Fleck S, Kalbe E, Rogers K, Hassenstab J, Brand M, et al. Introducing MASC: A movie for the assessment of social cognition. *Journal of Autism and Developmental Disorders*. 2006;36(5):623–36. Available from: <https://doi.org/10.1007/s10803-006-0107-0>
43. Goueli T, Nasreldin M, Madbouly N, Dziobek I, Farouk M. Social cognition in adolescent females with borderline personality traits. *Psychol Psychother*. 2020 Dec 1 [cited 2022 Feb 9];93(4):739–53. Available from: <https://pubmed.ncbi.nlm.nih.gov/31692159/>
44. Preißler S, Dziobek I, Ritter K, Heekeren H, Roepke S. Social Cognition in Borderline Personality Disorder: Evidence for Disturbed Recognition of the Emotions, Thoughts, and Intentions of others. Vol. 4, *Frontiers in Behavioral Neuroscience*. 2010. Available from: <https://www.frontiersin.org/article/10.3389/fnbeh.2010.00182>
45. Somma A, Ferrara M, Terrinoni A, Frau C, Ardizzone I, Sharp C, et al. Hypermentalizing as a marker of borderline personality disorder in Italian adolescents: a cross-cultural replication of Sharp and colleagues' (2011) findings. *Borderline Personal Disord Emot Dysregul*. 2019 Apr 10 [cited 2022 Feb 9];6(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/31007932/>
46. Jeung H, Herpertz SC. Impairments of interpersonal functioning. *Psychopathology*. 2014;47(4):220–34. Available from: <https://doi.org/10.1159/000357191>
47. Bertsch K, Hillmann K, Herpertz SC. Behavioral and Neurobiological Correlates of Disturbed Emotion Processing in Borderline Personality Disorder. *Psychopathology*. 2018;51(2):76–82. Available from: <https://www.karger.com/DOI/10.1159/000487363>
48. Baron-Cohen S, Jolliffe T, Mortimore C, Robertson M. Another advanced test of theory of mind: Evidence from very high functioning adults with autism or Asperger syndrome. *Journal of Child Psychology and Psychiatry and Allied Disciplines*. 1997;
49. Nilsen ES, Glenwright M, Huyder V. Children and adults understand that verbal irony interpretation depends on listener knowledge. *Journal of Cognition and Development*. 2011;12(3):374–409. Available from: <https://doi.org/10.1080/15248372.2010.544693> TS - CrossRef

50. Winner E, Gardner H. Metaphor and irony: Two levels of understanding. In: Ortony A, editor. *Metaphor and Thought*. 2nd ed. Cambridge: Cambridge University Press; 2012. p. 425–44. Available from: <https://doi.org/10.1017/CBO9781139173865.021>
51. Luyten P, de Meulemeester C, Fonagy P. The Self–Other Distinction in Psychopathology: Recent Developments from a Mentalizing Perspective. In: Gilead M, Ochsner KN, editors. *The Neural Basis of Mentalizing*. Cham: Springer International Publishing; 2021. p. 659–80. Available from: https://doi.org/10.1007/978-3-030-51890-5_34
52. de Meulemeester C, Lowyck B, Luyten P. The role of impairments in self–other distinction in borderline personality disorder: A narrative review of recent evidence. *Neuroscience & Biobehavioral Reviews*. 2021;127:242–54. Available from: <https://www.sciencedirect.com/science/article/pii/S0149763421001755>
53. de Meulemeester C, Lowyck B, Panagiotopoulou E, Fotopoulou A, Luyten P. Self–other distinction and borderline personality disorder features: Evidence for egocentric and altercentric bias in a self–other facial morphing task. *Personality Disorders: Theory, Research, and Treatment*. 2020 [cited 2022 Feb 17]; Available from: [/doiLanding?doi=10.1037/per0000415](https://doi.org/10.1037/per0000415)
54. Barnow S, Stopsack M, Grabe HJ, Meinke C, Spitzer C, Kronmüller K, et al. Interpersonal evaluation bias in borderline personality disorder. *Behaviour Research and Therapy*. 2009;47(5):359–65. Available from: <https://doi.org/10.1016/j.brat.2009.02.003>
55. Fertuck EA, Grinband J, Stanley B. Facial Trust Appraisal Negatively Biased in Borderline Personality Disorder. *Psychiatry Res*. 2013 May 30 [cited 2021 Sep 11];207(3):195. Available from: [/pmc/articles/PMC3654093/](https://pubmed.ncbi.nlm.nih.gov/2467692/)
56. Fonagy P, Luyten P, Allison E, Campbell C. What we have changed our minds about. *Borderline Personality Disorder and Emotion Dysregulation*. 2017;4(1):9. Available from: <https://doi.org/10.1186/s40479-017-0062-8>
57. Allen JG, Fonagy P, editors. *Handbook of mentalization-based treatment*. Chichester: John Wiley & Sons Ltd; 2006. 340 p. Available from: <http://gso.gbv.de/DB=2.1/PPNSET?PPN=1011605740>
58. Catalan A, Gonzalez de Artaza M, Bustamante S, Orgaz P, Osa L, Angosto V, et al. Differences in Facial Emotion Recognition between First Episode Psychosis, Borderline Personality Disorder and Healthy Controls. *PLoS One*. 2016 Jul 28;11(7):e0160056–e0160056. Available from: <https://pubmed.ncbi.nlm.nih.gov/27467692>
59. Domes G, Schulze L, Herpertz SC. Emotion recognition in borderline personality disorder—a review of the literature. *Journal of Personality Disorders*. 2009;23(1):6–19. Available from: <https://doi.org/10.1521/pedi.2009.23.1.6>
60. Kaiser D, Jacob GA, van Zutphen L, Siep N, Sprenger A, Tuschen-Caffier B, et al. Biased attention to facial expressions of ambiguous emotions in borderline personality disorder: An eye-tracking study. *Journal of Personality Disorders*. 2019;33(5):671–90. Available from: https://doi.org/10.1521/pedi_2019_33_363
61. Kieckhafer C, Felsenheimer AK, Rapp AM. A New Test for Irony Detection. *Frontiers in Psychiatry*. 2019;10:99. Available from: <https://doi.org/10.3389/fpsy.2019.00028>
62. Fydrich T, Renneberg B, Schmitz B, Wittchen HU. *Strukturiertes Klinisches Interview für DSM-IV Achse II: Persönlichkeitsstörungen*. Göttingen: Hogrefe; 1997. Available from: <http://www.sgip.org/diagnos/SKID.htm>
63. Wittchen HU, Zaudig M, Fydrich T. *SKID. Strukturiertes Klinisches Interview für DSM-IV Achse I*. Göttingen: Hogrefe; 1997.
64. Lehl S, Triebig G, Fischer B. Multiple choice vocabulary test MWT as a valid and short test to estimate premorbid intelligence. *Acta Neurologica Scandinavica*. 1995;91(5):335–45.
65. Bohus M, Kleindienst N, Limberger MF, Stieglitz RD, Domsalla M, Chapman AL, et al. The short version of the Borderline Symptom List (BSL-23): Development and initial data on psychometric properties. *Psychopathology*. 2009;42(1):32–9.
66. Klein C, Andresen B, Jahn T. Psychometric assessment of the schizotypal personality according to DSM-III-R criteria. Psychometric properties of an authorized German translation of Raine’s “Schizotypal Personality Questionnaire” (SPQ). *Diagnostica*. 1997;43(4):347–69.
67. Raine A. The SPQ. *Schizophr Bull*. 1991;17(4):555–64.

68. Davis MH. Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*. 1983;44(1):113–26.
69. Paulus C. Der Saarbrücker Persönlichkeitsfragebogen (SPF-IRI) zur Messung von Empathie. 2009; Available from: <http://hdl.handle.net/20.500.11780/3343>
70. Giora R. On Irony and Negation. *Discourse Processes*. 1995;19(2):239–64.
71. Kreuz RJ, Link KE. Asymmetries in the use of verbal irony. *Journal of Language and Social Psychology*. 2002;21(2):127–143 + 191.
72. Clark HH, Gerrig RJ. On the pretense theory of irony. *Journal of Experimental Psychology: General*. 1984;113(1):121–6. Available from: <https://doi.org/10.1037/0096-3445.113.1.121>
73. Gibbs R. Irony in Talk Among Friends. *Metaphor and Symbol*. 2000;15(1):5–27. Available from: <https://doi.org/10.1080/10926488.2000.9678862> TS - CrossRef
74. Herpertz SC, Huprich SK, Bohus M, Chanen A, Goodman M, Mehlum L, et al. The Challenge of Transforming the Diagnostic System of Personality Disorders. https://doi.org/10.1521/pedi_2017_31_338. 2017 Oct 12 [cited 2022 Apr 22];31(5):577–89. Available from: https://guilfordjournals.com/doi/abs/10.1521/pedi_2017_31_338
75. Felsenheimer A, Kieckhafer C, Rapp AM. Familiarity, empathy and comprehension of metaphors in patients with borderline personality disorder. *Psychiatry Research*. 2020;291:113152. Available from: <https://www.sciencedirect.com/science/article/pii/S0165178120301293>
76. Rapp A, Felsenheimer AK, Langohr K, Klupp M. The comprehension of familiar and novel metaphoric meanings in schizophrenia: A pilot study. *Frontiers in Psychology*. 2018;8(JAN):2251. Available from: <https://doi.org/10.3389/fpsyg.2017.02251>
77. Searle JR. Indirect Speech Acts. In: Cole P, Morgan JL, editors. *Speech Acts*. Leiden, Niederlande: Brill; 1975. p. 59–82. Available from: <https://brill.com/view/book/edcoll/9789004368811/BP000004.xml>
78. Tobin V, Israel M. Irony as a viewpoint phenomenon. In: Dancygier B, Sweetser E, editors. *Viewpoint in Language*. Cambridge: Cambridge University Press; 2012 [cited 2022 Feb 8]. p. 25–46. Available from: https://www.cambridge.org/core/product/identifier/9781139084727%23c01783-200/type/book_part
79. Orme W, Bowersox L, Vanwoerden S, Fonagy P, Sharp C. The relation between epistemic trust and borderline pathology in an adolescent inpatient sample. *Borderline Personality Disorder and Emotion Dysregulation*. 2019;6(1):13. Available from: <https://doi.org/10.1186/s40479-019-0110-7>
80. Harari H, Shamay-Tsoory SG, Ravid M, Levkovitz Y. Double dissociation between cognitive and affective empathy in borderline personality disorder. *Psychiatry Research*. 2010;175(3):277–9. Available from: <http://www.sciencedirect.com/science/article/pii/S0165178109000936>
81. New AS, aan het Rot M, Ripoll LH, Mercedes Perez-Rodriguez M, Lazarus S, Zipursky E, et al. Empathy and alexithymia in borderline personality disorder: Clinical and laboratory measures. *Journal of Personality Disorders*. 2012;26(5):660–75.
82. Petersen R, Brakoulias V, Langdon R. An experimental investigation of mentalization ability in borderline personality disorder. *Comprehensive Psychiatry*. 2016;64:12–21. Available from: <http://www.sciencedirect.com/science/article/pii/S0010440X15300237>
83. McLaren V, Gallagher M, Hopwood CJ, Sharp C. Hypermentalizing and Borderline Personality Disorder: A Meta-Analytic Review. *American Journal of Psychotherapy*. 2022 Jan 31;appi.psychotherapy.20210018. Available from: <https://doi.org/10.1176/appi.psychotherapy.20210018>
84. Dews S, Winner E. Muting the Meaning A Social Function of Irony. *Metaphor and Symbolic Activity*. 1995;10(1):3–19.
85. Pexman PM, Olineck KM. Does Sarcasm Always Sting? Investigating the Impact of Ironic Insults and Ironic Compliments. *Discourse Processes*. 2002;33(3):199–217. Available from: file:///C:/Users/Anne/Documents/Citavi 5/Projects/Doktorarbeit/Citavi Attachments/Pexman, Olineck 2002 - Does sarcasm always sting.pdf TS - EndNote Tagged Import Format M4 - Citavi

86. Weinbrecht A, Roepke S, Renneberg B. Fear of positive evaluation in borderline personality disorder. Taubner S, editor. PLOS ONE. 2020 Aug 20 [cited 2022 Feb 8];15(8):e0237944. Available from: <https://dx.plos.org/10.1371/journal.pone.0237944>
87. Reichenberger J, Eibl JJ, Pfaltz M, Wilhelm FH, Voderholzer U, Hillert A, et al. Don't Praise Me, Don't Chase Me: Emotional Reactivity to Positive and Negative Social-Evaluative Videos in Patients With Borderline Personality Disorder. *Journal of Personality Disorders*. 2016 Feb 4;31(1):75–89. Available from: https://doi.org/10.1521/pedi_2016_30_238
88. Winter D, Herbert C, Koplin K, Schmahl C, Bohus M, Lis S. Negative Evaluation Bias for Positive Self-Referential Information in Borderline Personality Disorder. PLOS ONE. 2015 Jan 22 [cited 2022 Feb 21];10(1):e0117083. Available from: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0117083>
89. Wiesenfeller J, Flasbeck V, Brown EC, Brüne M. Approach and Avoidance Behavior in Female Patients With Borderline Personality Disorder. Vol. 14, *Frontiers in Behavioral Neuroscience*. 2020. Available from: <https://www.frontiersin.org/article/10.3389/fnbeh.2020.588874>
90. Schneider I, Boll S, Volman I, Roelofs K, Spohn A, Herpertz SC, et al. Oxytocin Normalizes Approach–Avoidance Behavior in Women With Borderline Personality Disorder. Vol. 11, *Frontiers in Psychiatry*. 2020. Available from: <https://www.frontiersin.org/article/10.3389/fpsy.2020.00120>
91. Bateman A, Fonagy P. *Mentalization-Based Treatment for Personality Disorders*. First edit. Mentalization-Based Treatment for Personality Disorders. Oxford: Oxford University Press; 2016. XIV, [2], 468 stron TS-WorldCat T4-A practic. Available from: <http://www.worldcat.org/oclc/1066109332>
92. Foxhall M, Hamilton-Giachrisis C, Button K. The link between rejection sensitivity and borderline personality disorder: A systematic review and meta-analysis. *British Journal of Clinical Psychology*. 2019 Sep 1;58(3):289–326. Available from: <https://doi.org/10.1111/bjc.12216>
93. Vestergaard M, Kongerslev MT, Thomsen MS, Mathiesen BB, Harmer CJ, Simonsen E, et al. Women With Borderline Personality Disorder Show Reduced Identification of Emotional Facial Expressions and a Heightened Negativity Bias. https://doi.org/10.1521/pedi_2019_33_409. 2020 Oct 19 [cited 2022 Feb 8];34(5):677–98. Available from: https://guilfordjournals.com/doi/abs/10.1521/pedi_2019_33_409
94. Veague HB, Hooley JM. Enhanced sensitivity and response bias for male anger in women with borderline personality disorder. *Psychiatry Res*. 2014 Mar 30 [cited 2022 Feb 8];215(3):687–93. Available from: <https://pubmed.ncbi.nlm.nih.gov/24485062/>
95. Minzenberg MJ, Poole JH, Vinogradov S. Social-emotion recognition in borderline personality disorder. *Comprehensive Psychiatry*. 2006 Nov 1;47(6):468–74. Available from: <http://www.sciencedirect.com/science/article/pii/S0010440X06000502>
96. Niedtfeld I, Defiebre N, Regenbogen C, Mier D, Fenske S, Kirsch P, et al. Facing the Problem: Impaired Emotion Recognition During Multimodal Social Information Processing in Borderline Personality Disorder. *J Pers Disord*. 2017 Apr 1 [cited 2022 Feb 21];31(2):273–88. Available from: <https://pubmed.ncbi.nlm.nih.gov/27064850/>
97. Fenske S, Lis S, Liebke L, Niedtfeld I, Kirsch P, Mier D. Emotion recognition in borderline personality disorder: effects of emotional information on negative bias. *Borderline Personal Disord Emot Dysregul*. 2015 Jun 26 [cited 2022 Feb 8];2(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/26401312/>
98. Kobeleva X, Seidel EM, Kohler C, Schneider F, Habel U, Derntl B. Dissociation of explicit and implicit measures of the behavioral inhibition and activation system in borderline personality disorder. *Psychiatry Research*. 2014 Aug 15;218(1–2):134–42.
99. Franzen N, Hagenhoff M, Baer N, Schmidt A, Mier D, Sammer G, et al. Superior “theory of mind” in borderline personality disorder: an analysis of interaction behavior in a virtual trust game. *Psychiatry Res*. 2011 May 15 [cited 2022 Feb 21];187(1–2):224–33. Available from: <https://pubmed.ncbi.nlm.nih.gov/21129781/>
100. Torgersen S, Kringlen E, Cramer V. The prevalence of personality disorders in a community sample. *The Science of Mental Health: Volume 7: Personality and Personality Disorder*. 2013;58(6):132–8.

101. Ullrich S, Coid J. The age distribution of self-reported personality disorder traits in a household population. *Journal of Personality Disorders*. 2009;23(2):187–200. Available from: <https://doi.org/10.1521/pedi.2009.23.2.187>
102. Linehan M. Cognitive-behavioral treatment of borderline personality disorder. *Diagnosis and treatment of mental disorders*. New York: Guilford Press; 1993. 558 p. Available from: <http://gso.gbv.de/DB=2.1/PPNSET?PPN=686037901>
103. Colston HL. Irony and Sarcasm. In: Attardo S, editor. *The Routledge Handbook of Language and Humor*. Taylor and Francis; 2017. p. 234–49.
104. Langdon R, Coltheart M. Recognition of metaphor and irony in young adults: the impact of schizotypal personality traits. *Psychiatry Research*. 2004;125(1):9–20. Available from: <https://www.sciencedirect.com/science/article/pii/S0165178103002828>
105. Rapp AM, Langohr K, Mutschler DE, Wild B. Irony and Proverb Comprehension in Schizophrenia: Do Female Patients “Dislike” Ironic Remarks? *Schizophrenia Research and Treatment*. 2014;
106. Booth WC. *A rhetoric of irony*. Chicago, Ill: The Univ. of Chicago Press; 2007.
107. Jensen TW, Høgenhaug SS, Kjølbye M, Bloch MS. Mentalizing Bodies: Explicit Mentalizing Without Words in Psychotherapy. Vol. 12, *Frontiers in Psychology*. 2021. Available from: <https://www.frontiersin.org/article/10.3389/fpsyg.2021.577702>
108. Fuchs T. The phenomenology and development of social perspectives. *Phenomenology and the Cognitive Sciences*. 2013;12(4):655–83. Available from: <https://doi.org/10.1007/s11097-012-9267-x>
109. Gallese V. Bodily selves in relation: embodied simulation as second-person perspective on intersubjectivity. *Philos Trans R Soc Lond B Biol Sci*. 2014 Apr 28;369(1644):20130177. Available from: <https://pubmed.ncbi.nlm.nih.gov/24778374>

Figures

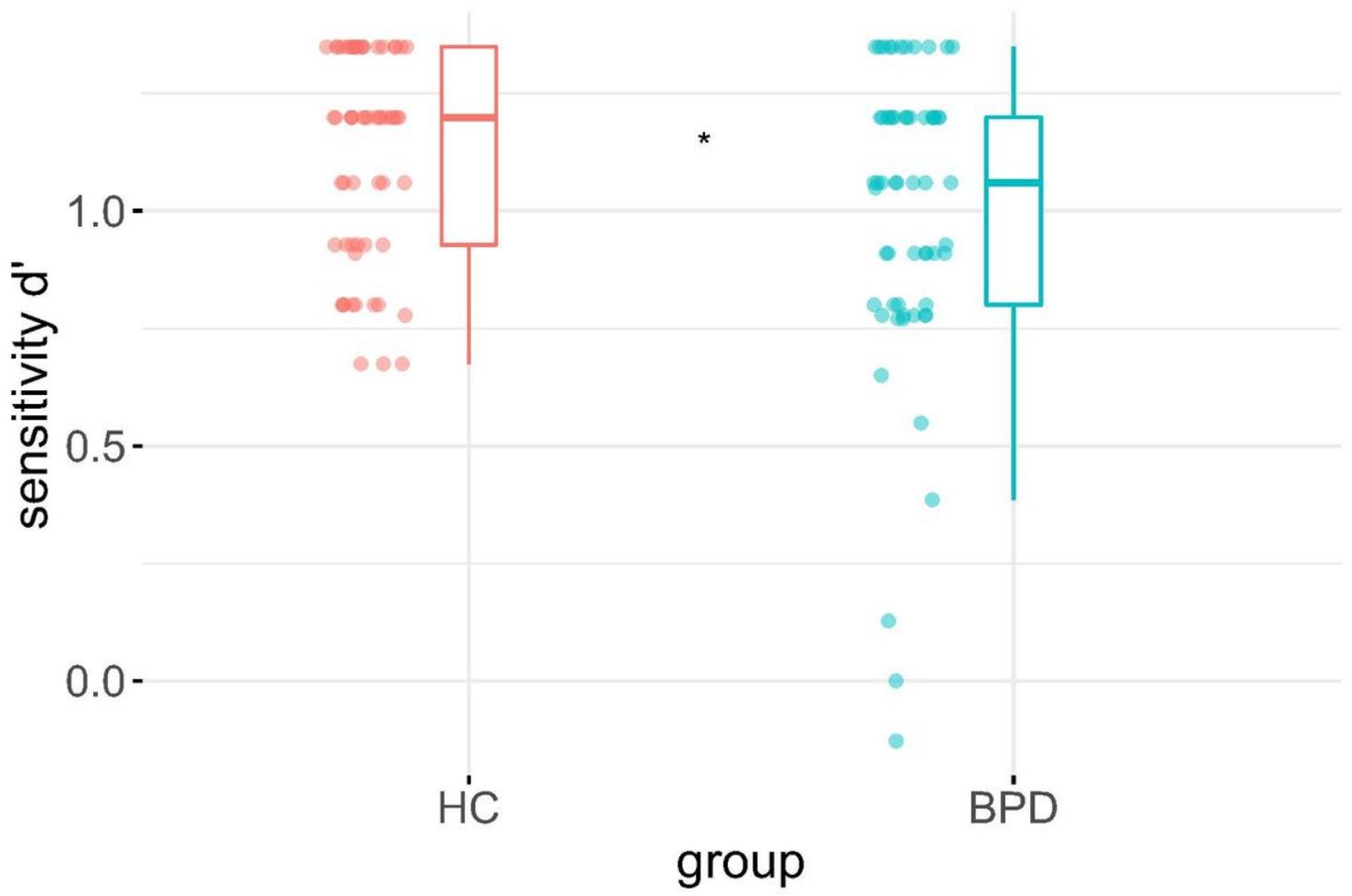


Figure 1

Sensitivity (d' prime) values for BPD and HC groups.

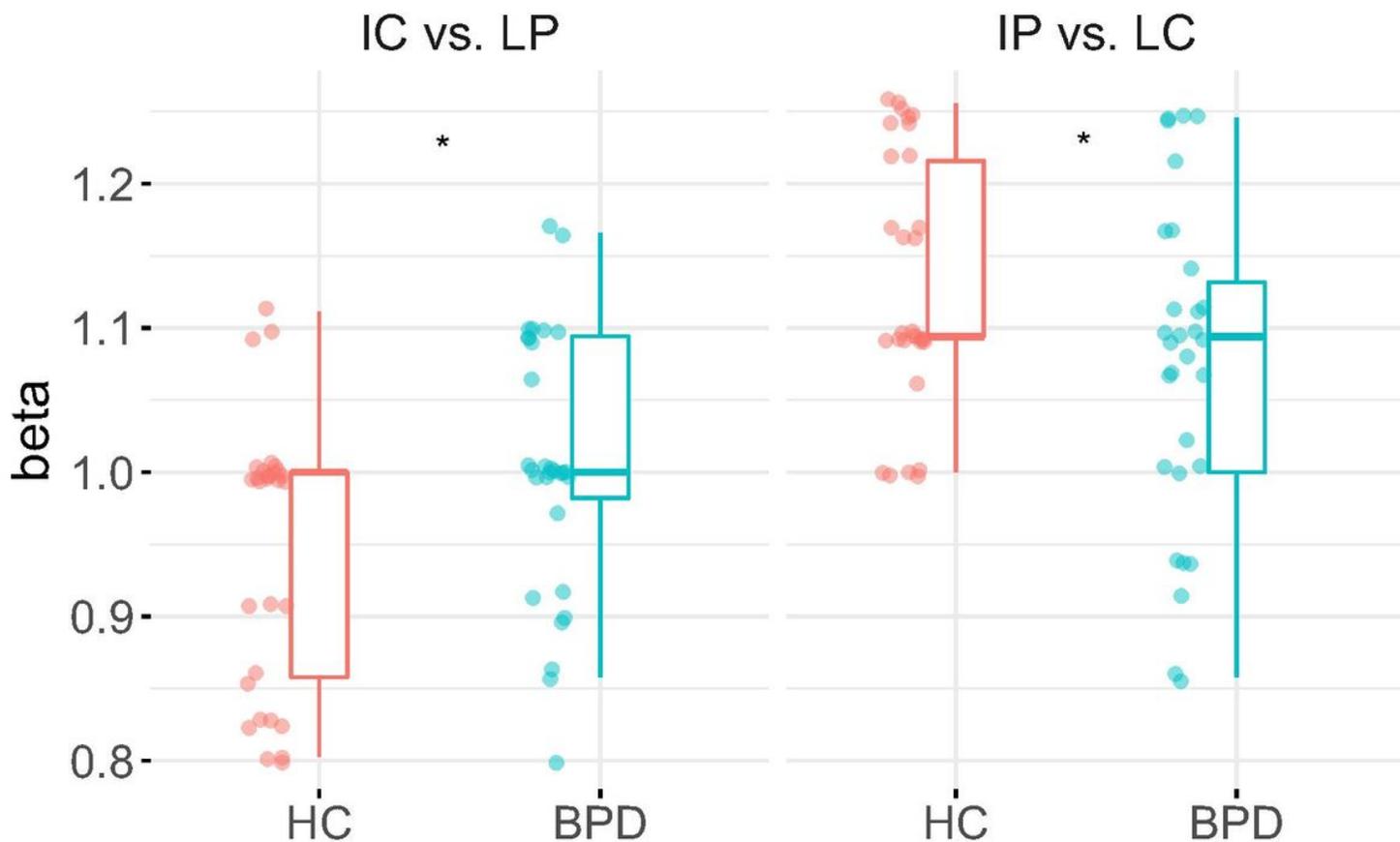


Figure 2

Response bias for the comparisons of ironic criticism (IC) with literal praise (LP) and ironic praise (IP) with literal criticism (LC). An unbiased β corresponds to 1, an β approaching zero a tendency to choose ironic, a β increasing over one a tendency to choose literal as response.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [AdditionalFile1.png](#)
- [AdditionalFile2.png](#)
- [Supplementarymaterial.docx](#)