Title: Characterization and Interrelationships of Theory of Mind, Socially Competitive Emotions and Affective Empathy in Bipolar Disorder

Authors: Lemvigh, C. K.[1,2,3], Karantonis, J.[3,4], Furlong, L. S.[3], Carruthers, S. P.[4], Pantelis, C.[1,3,5], Rossell, S. L.[4,6] & Van Rheenen, T. E.[3,4]*

Affiliations

1. Center for Clinical Intervention and Neuropsychiatric Schizophrenia Research (CINS) and Center for Neuropsychiatric Schizophrenia Research (CNSR), Mental Health Center Glostrup, Glostrup, Denmark
2. Department of Psychology, Faculty of Social Sciences, University of Copenhagen, Denmark
3. Melbourne Neuropsychiatry Centre, Department of Psychiatry, University of Melbourne and Melbourne Health, Melbourne, Australia
4. Centre for Mental Health, Faculty of Health, Arts and Design, Swinburne University of Technology, Victoria, 3122, Australia
5. Florey Institute of Neuroscience and Mental Health, Melbourne, Australia
6. Department of Psychiatry, St Vincent's Hospital, VIC, Australia

***Correspondence to**: Tamsyn E. Van Rheenen, Level 3, Alan Gilbert Building, 161 Barry Street, Carlton, Victoria, 3053, Australia, email: tamsyn.van@unimelb.edu.au

ABSTRACT

Evidence shows impaired Theory of Mind (ToM) in patients with bipolar disorder (BD), yet research examining its cognitive and affective components simultaneously is sparse. Moreover, recognition of socially competitive 'fortune of others' emotions (e.g. envy/gloat) may be related to ToM, but has not been assessed in BD. Finally, if and how ToM and 'fortune of others' emotions relate to affective empathy in BD is currently unclear. In this study, 64 BD patients and 34 healthy controls completed the Yoni task, a visual task assessing first- and second-order cognitive and affective ToM as well as 'fortune of others' emotions. The Toronto Empathy Questionnaire was used to assess self-reported affective empathy. Patients with BD showed no deficits in cognitive and affective ToM or recognition of 'fortune of others' emotions. The ability to infer 'fortune of others' emotions correlated with several ToM measures, indicating that these functions are part of the same system. Patients with BD reported similar levels of affective empathy to healthy controls, and this was not related to ToM or 'fortune of others' emotions, suggesting that affective empathy represents a separate social domain. These findings highlight areas of spared social functioning in BD, which may be utilized in therapeutic strategies.

Keywords: Social cognition, mentalizing , empathic abilities, mood disorder

INTRODUCTION

A substantial proportion of individuals with bipolar disorder (BD) experience cognitive deficits, which are known to contribute to real-world functional disability (Bowie et al., 2010; Van Rheenen et al., 2020a). These deficits occur in several domains, including social cognition (Miskowiak et al., 2019; Van Rheenen et al., 2019). In BD, social cognition may be partially independent of general cognition, with some evidence indicating that social cognition moderates the relationship between general cognition and functioning in patients with the illness (Ospina et al., 2018; Van Rheenen et al., 2014). These data thus highlight social cognitive deficits as an important component of BD's neuropsychological profile. Relevantly, current evidence suggests that social cognitive deficits in BD are quite subtle, particularly during the euthymic phase (Samamé et al., 2015; Van Rheenen and Rossell, 2014). Since it is known that subjective and objective reporting of cognitive functioning do not necessarily correspond (Burdick et al., 2005; Demant et al., 2015), social cognitive deficits may thus exist in patients with BD irrespective of self-report and may subtly contribute to the known difficulties many patients experience in terms of psychosocial functioning (Coryell et al., 1993; MacQueen et al., 2001).

Theory of Mind (ToM) is an important component of social cognition that refers to the ability to attribute mental states (beliefs, intentions and desires) to oneself and others (Povinelli and Preuss, 1995; Premack and Woodruff, 1978). It is a multifaceted construct consisting of several distinct components (Bora et al., 2016), including the capacity to a) understand the *knowledge and beliefs* of others (cognitive ToM), and b) understand the *emotions* of others (affective ToM) (Bodden et al., 2013; Bora et al., 2016; Shamay-Tsoory et al., 2007a). The ability to decode social cues in order to infer others' thoughts and emotions in social contexts is crucial in predicting the actions of others and facilitating prosocial behaviors (Van Rheenen et al., 2019). These skills are essential to social and functional success, which

are widely reported as lacking in BD even beyond the long-term resolution of clinical symptoms (Coryell et al., 1993; Dickerson et al., 2001; Greenberg et al., 2014; MacQueen et al., 2001; Pope et al., 2007).

Several existing studies indicate impairments of ToM in BD (Bora et al., 2016; Rossell and Van Rheenen, 2013; Samamé et al., 2012; Van Rheenen and Rossell, 2013), although there are exceptions (Caletti et al., 2013; Dalkner et al., 2019; Feyerabend et al., 2018; Haag et al., 2016; Purcell et al., 2013). These impairments have been reported in both cognitive and affective domains. However, the majority of the research has examined the two separately using different tasks, which has the potential to increase motivational/fatigue-related deficits that may confound findings. Studies examining cognitive and affective ToM in BD patients in the same task generally indicate impairments in the former, but not the latter (Andrews et al., 2016; Montag et al., 2010; Shamay-Tsoory et al., 2009; Wang et al., 2018).

The Yoni task was specifically designed to capture both cognitive and affective aspects of ToM in a single experiment (Shamay-Tsoory and Aharon-Peretz, 2007). An advantage of this visual task is that it requires minimal language and working memory demands and further enables the distinction between first-order (understanding what someone is thinking or feeling) and second-order ToM (understanding what someone thinks or feels about what someone else is thinking or feeling) (Shamay-Tsoory and Aharon-Peretz, 2007), thus providing the opportunity to deconstruct the multifaceted construct of ToM into specific components. Only two small studies have examined performance on the Yoni task in BD patients (Andrews et al., 2016; Wang et al., 2018). Both studies reported intact first-order ToM, and second-order affective ToM, whereas second-order cognitive ToM was significantly impaired (Andrews et al., 2016; Wang et al., 2018). Additionally, performance on the cognitive components of the Yoni task has been associated with hypomanic traits in males in the general population, with higher scores on the Hypomanic Personality Scale predicting lower ToM performance

(Terrien et al., 2014). Together these studies tentatively suggest that simple ToM, as assessed with the Yoni task, is spared in BD, and that the deficits observed in more complex ToM is due to difficulties in the ability to infer others' knowledge and beliefs, rather than their feeling states.

One version of the Yoni task includes an additional test assessing the ability to make inferences about socially competitive, so-called *'fortune of others' emotions*. These comprise 'gloating', 'envy' and 'identifying with someone', which are complex emotional states predominantly experienced in social situations. They may therefore be related to ToM (Shamay-Tsoory, 2008). Inferences about these complex emotions is thought to be mediated by the ventromedial prefrontal cortex, a brain region that is consistently implicated in ToM and also BD generally (Furlong et al., 2021; Phillips et al., 2008; Shamay-Tsoory et al., 2007b). Moreover, the ability to recognize these 'fortune of others' emotions has been found to correlate with ToM performance on the Yoni task specifically (Shamay-Tsoory, 2008; Shamay-Tsoory et al., 2007b). To the best of our knowledge, recognition of 'fortune of others' emotions has not been assessed in patients with BD. However, it is plausible that this subdomain is of relevance to BD given both its overlap with ToM and dependence on the ventromedial prefrontal cortex (Ha et al., 2009; Hiser and Koenigs, 2018).

Empathy, another important aspect of social cognition, can likewise be divided into cognitive and affective components (Dvash and Shamay-Tsoory, 2014; Kanske et al., 2016). Empathy is also related to psychosocial functioning and successful social communication and interactions (Cusi et al., 2010; Seidel et al., 2012). Despite sparse research attention, some preliminary work shows that BD patients have impaired cognitive empathy, yet intact or even superior affective empathy relative to controls (Bodnar & Rybakowski, 2017; Seidel et al., 2012; Shamay-Tsoory et al., 2009). This latter finding is unsurprising given that awareness of one's own emotions - a prerequisite for the embodiment of another's

- does not differ between BD patients and controls (Becerra et al., 2013; Van Rheenen et al., 2020b, 2015). Cognitive empathy is conceptually the same as ToM, in that it requires the capacity to *understand* another's thoughts, knowledge or emotions. In contrast, affective empathy refers to the ability to *sense or feel* other people's emotions (Dvash and Shamay-Tsoory, 2014; Shamay-Tsoory et al., 2009). Performance on the Yoni task has been found to correlate with responses on an affective empathy questionnaire in patients with acquired brain lesions (Shamay-Tsoory and Aharon-Peretz, 2007). While this correlation suggests overlap between ToM and affective empathy, other evidence points to a dissociation between these abilities (Kanske et al., 2016). The relationship between ToM and affective empathy in patients with BD is currently unclear. This is important to clarify however, because knowledge about whether there is a general system underlying social understanding in BD, as opposed to one that is modular, is relevant to the development of social cognitive treatments.

To improve our characterization of the social cognitive profile in BD, a better understanding of ToM, complex emotion recognition, affective empathy, and their inter-relationships is needed. Specifically, detailed knowledge about these components of social cognition in BD could increase our understanding of the underlying pathological processes and potentially pinpoint important targets for treatment. Thus, the aims of the current study were to examine; 1) first- and second-order cognitive and affective ToM; 2) the ability to infer complex 'fortune of others' emotions, and 3) affective empathy in patients with BD compared to controls. Based on the current literature, we hypothesized that first-order cognitive and affective ToM as well as second-order affective ToM would be spared, whereas second-order cognitive ToM would be impaired in BD patients. Moreover, we expected some impairments in recognition of 'fortune of others' emotions, but self-reported affective empathy levels similar to controls. A further exploratory question pertained to the presence and extent of inter-relationships between these sub-components of social cognition themselves, and also in relation to clinical measures in the BD group.

METHODS

The study was approved by the local Human Research Ethics Committee, and written informed consent was obtained from all participants.

Participants

A total of 64 stable outpatients with BD (BD I, $n = 60$; BD II, $n = 4$) and 34 healthy controls (HC) participated in the study (age 18-60). Participants were recruited as part of a larger neuroimaging study, using general advertisements, online websites and social media, with the BD patients also being recruited through community support groups and databases of participants held by the senior author. BD diagnoses were confirmed using the Mini International Neuropsychiatric Interview (MINI) (Lecrubier et al., 1997) according to DSM-IV-TR criteria (American Psychiatric Association, 2000). At the time of testing, no BD participant met criteria for a current mood episode. The majority of patients ($n = 52$) were considered affectively stable, as defined by scores of $\leq$12 on the Montgomery Asberg Depression Rating Scale (MADRS) (Davidson et al., 1986), and $\leq$8 on the Young Mania Rating Scale (YMRS) (Young et al., 1978); 45 patients also met strict criteria for euthymia defined as $\leq$8 on MADRS, and $\leq$8 on YMRS.

Exclusion criteria included: neurological/neurodegenerative disease, a history of severe head injury (loss of consciousness), hearing or visual impairments, pregnancy, language difficulties, a history of habitual drug use or dependence in the past three months prior to participation, and significant medication change in the two months prior to study inclusion. Further exclusion criteria for the HC group included: psychiatric illness (corroborated using MINI), family history of mood/psychiatric disorders, or current psychopharmacological medication/psychiatric care.

Measures

Descriptive measures: Information regarding medication use and dosage, age of onset, age at diagnosis, number of hospitalizations, psychosis history, and number of manic and depressive episodes were collected by self-report. Premorbid intelligence (IQ) was estimated using the Wechsler Test of Adult Reading (WTAR) (Wechsler, 2001) (see Table 1).

Theory of mind and inferring 'fortune of others' emotions were assessed using the Yoni task (Shamay-Tsoory, 2008). This is a computerized test programmed using E-prime (the task was downloaded from https://sites.google.com/edu.haifa.ac.il/sans/lab-tasks). Participants are shown a cartoon outline of a face (called Yoni) in the middle of the screen and asked to make judgements about his mental state based on verbal cues, facial expressions or eye gaze cues. Each corner of the screen contains a picture (either objects or faces) and the participant must point to the one Yoni is referring to, based on a sentence at the top of the screen (See Figure 1 and 2 in Shamay-Tsoory (2008)). The participants are instructed to respond as fast as possible.

The Yoni task measures *first and second-order ToM* in two blocks. The first-order ToM block comprises 12 cognitive trials (Yoni is thinking of _), 12 affective trials (Yoni loves _) and a control condition, including 8 physical items without the mentalizing component (Yoni is close to _). The second-order ToM block comprises 12 cognitive trials (Yoni is thinking about the flower that _ wants), 12 affective trials (Yoni loves the animal that _ does not love) and 6 physical control trials (Yoni has the toy that ___ has). In some trials, Yoni´s facial expression is neutral and in some it is emotionally loaded.

The additional *'fortune of others'* block is inbuilt into the Yoni task. Within this, there are three conditions in which participants are asked to identify the character that Yoni is gloating over (6 trials), envies (6 trials), or identifies with (6 trials), as well as a control condition requiring Yoni's facial

expression to be matched to other characters (physical #2, 6 trials). In all 'fortune of others' conditions, a combination of Yoni and the other characters´ facial expressions must be used to make the judgment.

Participants completed the task on a Lenovo computer (14″). Responses were made with a mouse click and the cursor was returned to the center of the screen after each response.

Affective empathy was measured using the Toronto Empathy Questionnaire (TEQ; Sprens et al., 2009). The TEQ is a short, self-report questionnaire with good psychometric properties that conceptualizes empathy as a primarily affective process (Sprens et al., 2009). It entails 16 statements, which participants are asked to rate how frequently they feel or act in the manner described. Responses are scored as "Never", "Rarely", "Sometimes", "Often", and "Always". Half of the items are positively worded, and half are negatively worded with reverse scoring. Scores are summarized into a total TEQ score, with higher scores indicative of higher affective empathy levels.

Statistical analyses

All statistical analyses were performed using SPSS (version 25.0, SPSS Inc.). Initial data checking indicated that YMRS, MADRS, WTAR and Yoni data were not normally distributed according to the Kolmogorov–Smirnov test, thus non-parametric tests were applied. Group differences in age, sex, estimated premorbid IQ (WTAR), and level of functioning were analyzed using *t*-tests, Chi-square tests and Mann-Whitney U. Given a significant group difference in WTAR scores ($U = 557.0$, $p < .001$, $r = -0.35$), which could affect the measures of interest given evidence of an association between ToM and cognitive ability (Olley et al., 2005), we further assessed whether the WTAR was associated with the Yoni data or affective empathy scores using Spearman's Rho correlations. As no significant correlations

were evident between these data (all r's < .21, all p's > .038), WTAR scores were not covaried in subsequent analyses.

Initially, group differences between BD patients and controls on the Yoni task and TEQ were examined using the Mann-Whitney U test. Effect sizes were calculated as $r = Z/\sqrt{N}$, with r values of 0.1, 0.3, and 0.5 indicative of small, medium and large effects respectively (Fritz et al., 2012). As a precautionary measure, one-sample t-tests were used to identify if there were any ceiling effects in the accuracy data of all conditions of the Yoni task. The TEQ total score was normally distributed and examined using an independent t-test and Cohen's d. Intercorrelations between the Yoni data and the TEQ were examined using Spearman's Rho in the full sample and also in the BD group itself. Further correlations examining the association between the variables of interest and clinical variables were conducted. Finally, effects of medications were explored. These latter two analyses were conducted in the BD sample only.

Re-analysis of all statistical tests comparing only those BD patients that were strictly euthymic to HCs was also conducted to examine the effect of current mood symptoms on the results. As the results of the re-analyses remained unchanged when case-control analyses were restricted to euthymic BD patients, for brevity only the analyses in the full BD sample are presented below (see supplementary material for the average Yoni performance in euthymic patients (**Table S1**)).

For each type of analysis, Bonferroni correction was applied to the alpha level to account for multiple comparisons (see supplementary material for more details regarding the number of tests used to calculate the Bonferroni corrected alpha value for each type of analysis). We also applied an alpha of p < .01 to verify our findings using a less stringent statistical correction method.

RESULTS

Demographic and clinical variables are presented in **Table 1**. The BD and HC groups were well matched, with no significant differences in age ($t(1, 95) = 0.529$, $p = .598$) or sex (χ^2 (1) = 0.492, p = .483).

[INSERT TABLE 1 HERE]

Yoni task and TEQ

Mean accuracy and reaction time (RT) data from all conditions within all blocks of the Yoni task and the TEQ are presented in **Table 2**. While no ceiling effects were evident (all ceiling effect results survived the Bonferroni-adjusted alpha level of $p < .005$; all t's > -3.08, all p's < .003), no results from the group difference analyses were significant using the Bonferroni-adjusted alpha level of $p < .002$ (.05/21) or even a less stringent alpha value of $p < .01$. Details are provided below.

First-order ToM: There were no significant differences between BD patients and HCs in accuracy scores in the cognitive ($U = 938.5$ $p = .112$, $r = -0.16$), affective ($U = 1048.5$, $p = .628$, $r = -0.05$) or physical ($U = 965.5$, $p = .210$, $r = -0.13$) conditions. No RT differences were evident in the affective ($U = 912.0$, $p = .189$, $r = -0.13$), or physical condition ($U = 894.0$, $p = .148$, $r = -0.15$). We did observe a significant difference in RT in the cognitive condition, with patients responding faster than controls using a less stringent alpha value of $p < .01$, however this did not survive Bonferroni correction ($U = 702.0$, $p = .004$, $r = -0.29$). Effect sizes were in the small range.

Second-order ToM: No significant accuracy differences were observed between BD patients and HCs in the affective ($U = 889.5$, $p = .131$ $r = -0.15$), cognitive ($U = 837.5$, $p = .058$, $r = -0.19$) or physical ($U = 1034.0$, $p = .651$ $r = -0.05$) conditions. There were no significant RT differences in any condition (cognitive: $U = 1012.0$, $p = .571$, $r = -0.06$, affective: $U = 1005.0$, $p = .536$, $r = -0.06$, physical: U =

997.0, p = .497, r = -0.07). Effect sizes were generally small, with the largest effect observed for second-order cognitive accuracy.

'Fortune of others' emotions: There were no significant group differences in accuracy in the gloating (U = 1059.5, p = .823, r = -0.02), envy *(*U = 901.5, p = .134, r = -0.15), identification (U = 1050.0, p = .750, r = -0.03) or physical #2 (U = 1065.0, p = .847, r = -0.02) conditions, nor were there differences in RT in any of these conditions (gloating: U = 1001.0 p = .516, r = -0.07, envy: U = 1040.0, p = .720, r = -0.04, identification: U = 977.0, p = .407, r = -0.08, physical #2: U = 1001.0, p = .516, r = -0.07). Effect sizes were in the small range, with the largest effect size observed for envy accuracy.

Affective empathy: There were no significant differences between BD and HCs on the TEQ total score (t(1, 92) = -0.883, p = .379, d = -0.19). The size of the case-control effect was small.

[INSERT TABLE 2 HERE]

Associations between variables

Correlations between Yoni task data and TEQ scores: In the full sample, envy accuracy correlated significantly with first-order affective (r_s(96) = .346, p = .0004), second-order cognitive, (r_s(96) = .372, p = .0002) and second-order affective accuracy (r_s(96) = .463, p <.0001), whereas gloat accuracy only correlated significantly with second-order affective accuracy (r_s(96) = .359, p = .0003), using a Bonferroni-adjusted alpha level of p < .007. Identification accuracy did not correlate with any of the ToM measures (all r's < .161, all p's > .112) (**Table 3**). When the sample was restricted to BD patients, envy

accuracy still correlated with second-order affective accuracy, while gloat accuracy correlated significantly with both second-order cognitive and second-order affective accuracy (**Supplementary Table S2**). Most of the RT variables from the ToM conditions correlated with the RT for 'fortune of others' emotion conditions both in the full sample (**Table 3**) and in BD patients only (**Supplementary Table S2**), indicating that fast responders were fast on most trials.

The TEQ total score did not correlate with Yoni task accuracy data from any of the conditions within first-order ToM, second-order ToM or 'fortune of others' emotions after Bonferroni corrections either in the full sample (all r's < .126 , all p's > .227) (**Table 3**) or in BD patients only (all r's < -.320 , all p's > .010) (**Supplementary Table S2**). The TEQ total score was correlated with RT from the envy condition ($r_s(92) = -.369$, $p = .0003$) in the full sample (**Table 3**).

[INSERT TABLE 3]

Associations between Yoni task data, TEQ scores, clinical measures and medications in the BD group: There were no significant associations between accuracy in Yoni task data and mood symptoms measured by the MADRS and YMRS in the BD patient group (all r's < -.184, all p's > .145). Performance on the Yoni task did not correlate with age of onset (all r's < -.195, all p's > .150), age at diagnosis (all r's < -.212, all p's > .107), number of hospitalizations (all r's < -.199, all p's > .127), number of manic episodes (all r's < .257, all p's > .045) or number of depressive episodes (all r's < .201, all p's > .123). Performance on the Yoni task did not correlate with mean lithium dose in the BD sample (all r's < .270, all p's > .236). Additionally, there were no significant differences (after Bonferroni correction) in Yoni performance between unmedicated patients, patients using typical antipsychotics, atypical

antipsychotics, mood stabilizers, antidepressants or a combination of medication types (all X^2's < 11.51, all *p's* > .042). Finally, there were no significant differences between patients with a history of psychosis (n=38) and patients without a history of psychosis (n=25) on either ToM or empathy measures (all U^2's > 410.0, all *p's* > .150).

DISCUSSION

The current study investigated ToM in a sample of patients with BD using a task designed to disentangle cognitive and affective ToM in the context of first-order and second-order ToM abilities. As expected, BD patients performed at the same level as HC's with regard to both cognitive and affective first-order ToM, consistent with the previous literature using the Yoni task in BD patients (Andrews et al., 2016; Wang et al., 2018). Also consistent with the two previous studies was the observation of no group differences in second-order ToM in the *affective* condition (Andrews et al., 2016; Wang et al., 2018). However, the previous two studies reported a significant deficit in second-order *cognitive* ToM, which we could not replicate in this sample using non-parametric statistics and Bonferroni corrections to account for multiple testing. It should be noted that we included a much larger sample compared to the two previous studies (64 patients with BD here, compared to 15 and 35 in previous work). Lack of power is therefore not a likely explanation for the discrepancy. Rather, it is possible that the discrepant findings may be explained by the use of different versions of the task, including differences in the number of trials for each condition. Alternatively, even though the two previous studies also included euthymic and clinically stable BD patients, it is possible that the differences between our findings could be explained by other sample characteristics, such as years of education, medication use or clinical course (e.g. history of mood episodes, hospitalizations or rapid cycling).

Overall, the findings of ToM performance equivalent to healthy controls is inconsistent with *some* of the broader literature showing ToM deficits in BD (although see Caletti et al., 2013; Dalkner et al., 2019; Feyerabend et al., 2018; Haag et al., 2016; Purcell et al., 2013 for exceptions). These discrepancies do not appear to be purely statistical given medium to large effect sizes in previous meta-analyses compared to the small effects seen in the current work (Bora et al., 2016; Samamé et al., 2012). Rather, they may be explained by the use of a visual task here, whereas the majority of past research has used verbal tasks. Some evidence shows that only verbal and not visual ToM may be an endophenotype for BD (Reynolds et al., 2014). Moreover, although there is evidence that social cognition in BD may be *partially* explained by general cognitive ability, the Yoni task was designed to have a very low memory load. In earlier work we showed moderate to large working memory (and other cognitive) deficits in several BD patients from this sample, even in the absence of ToM deficits here (Karantonis et al., 2020). Hence, it is unlikely that an absence of cognitive deficits can explain these negative findings. Instead, it is possible that previous ToM studies showing impairments with more cognitively taxing tasks were not entirely indexing ToM itself, but also general cognitive deficits. Finally, it is also possible that the Yoni task is not sensitive enough to elicit subtle ToM difficulties.

There were no group differences in recognition of 'fortune of others' emotions. To the best of our knowledge, the ability to infer these complex socially competitive emotional states has not previously been investigated in BD, and our findings suggest that this ability is not disproportionately affected in those with the disorder. In line with previous literature, recognition of envy and gloat correlated significantly with several ToM measures (Shamay-Tsoory, 2008; Shamay-Tsoory et al., 2007b). These findings suggest that the ability to understand socially competitive emotions is related to general ToM abilities in BD patients. Hence, the absence of group differences in the recognition of these emotions is not surprising given that ToM was not impaired in the BD group.

Previous studies in BD have reported impairments in cognitive empathy, whereas affective empathy may be spared or even increased especially during manic phases (Bodnar & Rybakowski, 2017; Cusi et al., 2010; Seidel et al., 2012; Shamay-Tsoory et al., 2009). In line with this, the BD patients in our sample reported comparable levels of affective empathy to HC's. We observed no associations between affective empathy and affective ToM or any other ToM measures, suggesting that ToM and affective empathy represent distinct concepts, consistent with previous findings that the abilities to mentalize and empathize are independent (Kanske et al., 2016; Stietz et al., 2019). An alternative explanation is that ToM was measured behaviorally, while affective empathy was examined using self-report, and these measures may not be equivalent (Dang et al., 2020; Murphy and Lilienfeld, 2019).

Finally, we did not observe any associations between ToM performance on the Yoni task and clinical variables of interest, i.e. YMRS, MADRS, estimated premorbid IQ, age of onset, age of diagnosis, number of hospitalizations, estimated lifetime manic and depressive episodes or mean lithium dose, consistent with previous findings using standard ToM tests in euthymic patients (Bora et al., 2005). However, the subsamples of patients on and off medications in our study were very small, thus caution should be taken in inferring that medications had no effect.

The present findings should be considered in the context of the strengths and limitations of the study. A key strength was that the use of the Yoni task allowed for the simultaneous assessment of both cognitive and affective ToM as well as first- and second-order levels of complexity. Thus, it provided a means of reducing the participant burden, and associated motivational/fatigue-related deficits, that may have arisen if using multiple different tasks to probe these subcomponents of ToM. This is also among the larger studies of ToM abilities in BD patients, giving the null findings reported here some credence. Indeed, the larger sample and minimal patient-control effect sizes compared to previous work on the topic, suggest that the negative findings observed here cannot merely be ascribed to power issues. A

caveat to this is that sample may not have been powered enough to detect *subtle* effects if they do exist, and the results should be interpreted with this in mind.

Other limitations include the smaller number of trials in the assessment of 'fortune of others' emotions compared to the ToM conditions of the Yoni task, which is a function of the task design. Some measures, including the TEQ and some clinical variables (e.g. past mood episodes, hospitalizations) were self-reported rather than clinician rated and should therefore be interpreted with caution. General cognitive performance was also not controlled in the analyses, thus we cannot be sure that the findings are independent of level of cognitive functioning. Nonetheless, we included the WTAR as an estimate of premorbid IQ, and despite group differences between BD patients and healthy controls, scores were not associated with the Yoni data. Finally, the use of a cross-sectional design did not allow us to examine how time-varying symptoms of BD influence ToM and recognition of 'fortune of others' emotions. To more thoroughly understand state and trait effects, a longitudinal study is needed to investigate if changes in performance on the Yoni task occur with changes in symptom levels in the same individuals. In our study, we observed similar findings when the sample was restricted to euthymic BD patients only, and none of the ToM measures correlated with mood symptoms. Thus, our findings do suggest that ToM and recognition of 'fortune of others' emotions in BD patients do not vary according to mood state. However, given the presence of null findings in this study in the context of previous research showing ToM deficits in euthymic BD patients, longitudinal work would help to clarify the nature of social cognitive functioning in BD even further.

Irrespective of these limitations, the current findings enhance our understanding of the social cognitive profile in BD by highlighting ToM and empathy as areas of potentially spared functioning. Given evidence that many patients with BD have experienced adversity during developmental periods in which these abilities develop, our findings suggest that intact ToM and empathy may be markers of

resilience in BD (Etain et al., 2008; Watson et al., 2014). Knowledge that ToM and empathy is intact is particularly relevant in light of an increased focus on the importance of considering cognitive strengths in the context of approaches to remediate cognitive deficits (Allott et al., 2020). This could be done by considering practical ways that a patients' knowledge of intact ToM and empathy could be utilized to reduce self-stigma and promote self-efficacy, improved well-being and functioning (Allott et al., 2020). Hence these findings are important for therapeutic interventions to support functional recovery in BD. In sum, we observed intact cognitive and affective ToM in BD patients using a visual task, no deficit in recognition of 'fortune of others' emotions, and levels of affective empathy comparable to HC's. There appears to be some overlap between ToM and recognition of 'fortune of others' emotions, whereas ToM and affective empathy seem to be unrelated. However, future studies should examine empathy behaviorally and include measures of cognitive empathy.

REFERENCES

Allott, K., Steele, P., Boyer, F., de Winter, A., Bryce, S., Alvarez-Jimenez, M., Phillips, L., 2020. Cognitive strengths-based assessment and intervention in first-episode psychosis: A complementary approach to addressing functional recovery? Clin. Psychol. Rev. 79, 101871. https://doi.org/10.1016/j.cpr.2020.101871

American Psychiatric Association, 2000. Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR). American Psychiatric Association.

Andrews, S.C., Enticott, P.G., Hoy, K.E., Thomson, R.H., Fitzgerald, P.B., 2016. Reduced mu suppression and altered motor resonance in euthymic bipolar disorder: Evidence for a dysfunctional mirror system? Soc. Neurosci. 11, 60–71. https://doi.org/10.1080/17470919.2015.1029140

Becerra, R., Cruise, K., Murray, G., Bassett, D., Harms, C., Allan, A., Hood, S., 2013. Emotion regulation in bipolar disorder: Are emotion regulation abilities less compromised in euthymic bipolar disorder than unipolar depressive or anxiety disorders? Open J. Psychiatry 3, 1–7. https://doi.org/10.4236/ojpsych.2013.34a001

Bodden, M.E., Kübler, D., Knake, S., Menzler, K., Heverhagen, J.T., Sommer, J., Kalbe, E., Krach, S., Dodel, R., 2013. Comparing the neural correlates of affective and cognitive theory of mind using fMRI: Involvement of the basal ganglia in affective theory of mind. Adv. Cogn. Psychol. 9, 32–43. https://doi.org/10.2478/v10053-008-0129-6

Bodnar, A., Rybakowski, J.K., 2017. Increased affective empathy in bipolar patients during a manic episode. Rev. Bras. Psiquiatr. 39, 342–345. https://doi.org/10.1590/1516-4446-2016-2101

Bora, E., Bartholomeusz, C., Pantelis, C., 2016. Meta-analysis of Theory of Mind (ToM) impairment in bipolar disorder. Psychol. Med. 46, 253–264. https://doi.org/10.1017/S0033291715001993

Bora, E., Vahip, S., Gonul, A.S., Akdeniz, F., Alkan, M., Ogut, M., Eryavuz, A., 2005. Evidence for theory of mind deficits in euthymic patients with bipolar disorder. Acta Psychiatr. Scand. 112, 110–116. https://doi.org/10.1111/j.1600-0447.2005.00570.x

Bowie, C.R., Depp, C., Mcgrath, J.A., Wolyniec, P., Mausbach, B.T., Thornquist, M.H., Luke, J., Patterson, T.L., Harvey, P.D., Pulver, A.E., 2010. Prediction of Real-World Functional Disability in Chronic Mental Disorders: A Comparison of Schizophrenia and Bipolar Disorder. Am J Psychiatry 167, 1116–1124.

Burdick, K.E., Endick, C.J., Goldberg, J.F., 2005. Assessing cognitive deficits in bipolar disorder: Are self-reports valid? Psychiatry Res. 136, 43–50. https://doi.org/10.1016/j.psychres.2004.12.009

Caletti, E., Paoli, R.A., Fiorentini, A., Cigliobianco, M., Zugno, E., Serati, M., Orsenigo, G., Grillo, P., Zago, S., Caldiroli, A., Prunas, C., Giusti, F., Consonni, D., Altamura, A.C., 2013. Neuropsychology, social cognition and global functioning among bipolar, schizophrenic patients and healthy controls: Preliminary data. Front. Hum. Neurosci. https://doi.org/10.3389/fnhum.2013.00661

Coryell, W., Scheftner, W., Keller, M., Endicott, J., Maser, J., Klerman, G.L., 1993. The enduring psychosocial consequences of mania and depression. Am. J. Psychiatry 150, 720–727.

Cusi, A., MacQueen, G.M., McKinnon, M.C., 2010. Altered self-report of empathic responding in patients with bipolar disorder. Psychiatry Res. 178, 354–358. https://doi.org/10.1016/j.psychres.2009.07.009

Dalkner, N., Bengesser, S.A., Birner, A., Fellendorf, F.T., Hamm, C., Platzer, M., Pilz, R., Queissner, R., Rieger, A., Weber, B., Kapfhammer, H.P., Weiss, E.M., Reininghaus, E.Z., 2019. The relationship between "Eyes Reading" ability and verbal memory in bipolar disorder. Psychiatry Res. 273, 42–51. https://doi.org/10.1016/j.psychres.2019.01.015

Dang, J., King, K.M., Inzlicht, M., 2020. Why Are Self-Report and Behavioral Measures Weakly Correlated? Trends Cogn. Sci. 24, 267–269. https://doi.org/10.1016/j.tics.2020.01.007

Davidson, J., Turnbull, C.D., Strickland, R., Miller, R., Graves, K., 1986. The Montgomery-Åsberg Depression Scale: reliability and validity. Acta Psychiatr. Scand. 73, 544–548. https://doi.org/10.1111/j.1600-0447.1986.tb02723.x

Demant, K.M., Vinberg, M., Kessing, L. V., Miskowiak, K.W., 2015. Assessment of subjective and objective cognitive function in bipolar disorder: Correlations, predictors and the relation to psychosocial function. Psychiatry Res. 229, 565–571. https://doi.org/10.1016/j.psychres.2015.05.022

Dickerson, F.B., Sommerville, J., Origoni, A.E., Ringel, N.B., Parente, F., 2001. Outpatients with schizophrenia and bipolar I disorder: Do they differ in their cognitive and social functioning? Psychiatry Res. 102, 21–27. https://doi.org/10.1016/S0165-1781(01)00247-5

Dvash, J., Shamay-Tsoory, S.G., 2014. Theory of mind and empathy as multidimensional constructs: Neurological foundations. Top. Lang. Disord. 34, 282–295.

https://doi.org/10.1097/TLD.0000000000000040

Etain, B., Henry, C., Bellivier, F., Mathieu, F., Leboyer, M., 2008. Beyond genetics: Childhood affective trauma in bipolar disorder. Bipolar Disord. 10, 867–876. https://doi.org/10.1111/j.1399-5618.2008.00635.x

Feyerabend, J., Lüttke, S., Grosse-Wentrup, F., Wolter, S., Hautzinger, M., Wolkenstein, L., 2018. Theory of mind in remitted bipolar disorder: Younger patients struggle in tasks of higher ecological validity. J. Affect. Disord. 231, 32–40. https://doi.org/10.1016/j.jad.2018.01.026

Fritz, C.O., Morris, P.E., Richler, J.J., 2012. Effect Size Estimates: Current Use, Calculations, and Interpretation. J. Exp. Psychol. Gen. 141, 2–18. https://doi.org/10.1037/a0024338

Furlong, L.S., Rossell, S.L., Caruana, G.F., Cropley, V.L., Hughes, M., Van Rheenen, T.E., 2021. The activity and connectivity of the facial emotion processing neural circuitry in bipolar disorder: a systematic review. J. Affect. Disord. 279, 518–548. https://doi.org/10.1016/j.jad.2020.10.038

Greenberg, S., Rosenblum, K.L., McInnis, M.G., Muzik, M., 2014. The role of social relationships in bipolar disorder: A review. Psychiatry Res. 219, 248–254. https://doi.org/10.1016/j.psychres.2014.05.047

Ha, T.H., Ha, K., Kim, J.H., Choi, J.E., 2009. Regional brain gray matter abnormalities in patients with bipolar II disorder: A comparison study with bipolar I patients and healthy controls. Neurosci. Lett. 456, 44–48. https://doi.org/10.1016/j.neulet.2009.03.077

Hiser, J., Koenigs, M., 2018. The Multifaceted Role of the Ventromedial Prefrontal Cortex in Emotion, Decision Making, Social Cognition, and Psychopathology. Biol. Psychiatry 83, 638–647. https://doi.org/10.1016/j.biopsych.2017.10.030

Haag, S., Haffner, P., Quinlivan, E., Brüne, M., Stamm, T., 2016. No differences in visual theory of mind abilities between euthymic bipolar patients and healthy controls. Int. J. Bipolar Disord. 4. https://doi.org/10.1186/s40345-016-0061-5

Kanske, P., Böckler, A., Trautwein, F.M., Lesemann, F.H.P., Singer, T., 2016. Are strong empathizers better mentalizers? Evidence for independence and interaction between the routes of social cognition. Soc. Cogn. Affect. Neurosci. 11, 1383–1392. https://doi.org/10.1093/scan/nsw052

Karantonis, J.A., Rossell, S.L., Carruthers, S.P., Sumner, P., Hughes, M., Green, M.J., Pantelis, C., Burdick, K.E., Cropley, V., Van Rheenen, T.E., 2020. Cognitive validation of cross-diagnostic cognitive subgroups on the schizophrenia-bipolar spectrum. J. Affect. Disord. 266, 710–721.

https://doi.org/10.1016/j.jad.2020.01.123

Lecrubier, Y., Sheehan, D. V., Weiller, E., Amorim, P., Bonora, I., Sheehan, K.H., Janavs, J., Dunbar, G.C., 1997. The Mini International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: Reliability and validity according to the CIDI. Eur. Psychiatry 12, 224–231. https://doi.org/10.1016/S0924-9338(97)83296-8

MacQueen, G.M., Trevor Young, L., Joffe, R.T., 2001. A review of psychosocial outcome in patients with bipolar disorder. Acta Psychiatr. Scand. 103, 163–170. https://doi.org/10.1034/j.1600-0447.2001.00059.x

Miskowiak, K.W., Seeberg, I., Kjaerstad, H.L., Burdick, K.E., Martinez-Aran, A., del Mar Bonnin, C., Bowie, C.R., Carvalho, A.F., Gallagher, P., Hasler, G., Lafer, B., López-Jaramillo, C., Sumiyoshi, T., McIntyre, R.S., Schaffer, A., Porter, R.J., Purdon, S., Torres, I.J., Yatham, L.N., Young, A.H., Kessing, L. V., Van Rheenen, T.E., Vieta, E., 2019. Affective cognition in bipolar disorder: A systematic review by the ISBD targeting cognition task force. Bipolar Disord. 21, 686–719. https://doi.org/10.1111/bdi.12834

Montag, C., Ehrlich, A., Neuhaus, K., Dziobek, I., Heekeren, H.R., Heinz, A., Gallinat, J., 2010. Theory of mind impairments in euthymic bipolar patients. J. Affect. Disord. 123, 264–269. https://doi.org/10.1016/j.jad.2009.08.017

Murphy, B.A., Lilienfeld, S.O., 2019. Are self-report cognitive empathy ratings valid proxies for cognitive empathy ability? Negligible meta-analytic relations with behavioral task performance. Psychol. Assess. 31, 1062–1072. https://doi.org/10.1037/pas0000732

Olley, A.L., Malhi, G.S., Bachelor, J., Cahill, C.M., Mitchell, P.B., Berk, M., 2005. Executive functioning and theory of mind in euthymic bipolar disorder. Bipolar Disord. 7, 43–52. https://doi.org/10.1111/j.1399-5618.2005.00254.x

Ospina, L.H., Nitzburg, G.C., Shanahan, M., Perez-Rodriguez, M.M., Larsen, E., Latifoglu, A., Burdick, K.E., 2018. Social cognition moderates the relationship between neurocognition and community functioning in bipolar disorder. J. Affect. Disord. 235, 7–14. https://doi.org/10.1016/j.jad.2018.03.013

Phillips, M.L., Ladouceur, C., Drevets, W.C., 2008. A neural model of voluntary and automatic emotion regulation: implications for understanding the pathophysiology and neurodevelopment of bipolar disorder. Mol. Psychiatry 13, 829–857. https://doi.org/10.1038/mp.2008.65.A

Pope, M., Dudley, R., Scott, J., 2007. Determinants of social functioning in bipolar disorder. Bipolar Disord. 9, 38–44. https://doi.org/10.1111/j.1399-5618.2007.00323.x

Povinelli, D.J., Preuss, T.M., 1995. Theory of mind: evolutionary history of a cognitive specialization. Trends Neurosci. 18, 418–424. https://doi.org/10.1016/0166-2236(95)93939-U

Premack, D., Woodruff, G., 1978. Does the chimpanzee have a theory of mind? Behav. Brain Sci. 4, 515–526.

Purcell, A.L., Phillips, M., Gruber, J., 2013. In your eyes: Does theory of mind predict impaired life functioning in bipolar disorder? J. Affect. Disord. 151, 1113–1119. https://doi.org/10.1016/j.jad.2013.06.051

Reynolds, M.T., Van Rheenen, T.E., Rossell, S.L., 2014. Theory of mind in first degree relatives of individuals with bipolar disorder. Psychiatry Res. 219, 400–402. https://doi.org/10.1016/j.psychres.2014.05.041

Rossell, S.L., Van Rheenen, T.E., 2013. Theory of mind performance using a story comprehension task in bipolar mania compared to schizophrenia and healthy controls. Cogn. Neuropsychiatry 18, 409–421. https://doi.org/10.1080/13546805.2012.725820

Samamé, C., Martino, D.J., Strejilevich, S.A., 2015. An individual task meta-analysis of social cognition in euthymic bipolar disorders. J. Affect. Disord. 173, 146–153. https://doi.org/10.1016/j.jad.2014.10.055

Samamé, C., Martino, D.J., Strejilevich, S.A., 2012. Social cognition in euthymic bipolar disorder: Systematic review and meta-analytic approach. Acta Psychiatr. Scand. 125, 266–280. https://doi.org/10.1111/j.1600-0447.2011.01808.x

Seidel, E.M., Habel, U., Finkelmeyer, A., Hasmann, A., Dobmeier, M., Derntl, B., 2012. Risk or resilience? Empathic abilities in patients with bipolar disorders and their first-degree relatives. J. Psychiatr. Res. 46, 382–388. https://doi.org/10.1016/j.jpsychires.2011.11.006

Shamay-Tsoory, S.G., 2008. Recognition of "Fortune of others" emotions in asperger syndrome and high functioning autism. J. Autism Dev. Disord. 38, 1451–1461. https://doi.org/10.1007/s10803-007-0515-9

Shamay-Tsoory, S.G., Shur, S., Barcai-Goodman, L., Medlovich, S., Harari, H., Levkovitz, Y., 2007a. Dissociation of cognitive from affective components of theory of mind in schizophrenia. Psychiatry Res. 149, 11–23. https://doi.org/10.1016/j.psychres.2005.10.018

Shamay-Tsoory, S.G., Tibi-Elhanany, Y., Aharon-Peretz, J., 2007b. The green-eyed monster and malicious joy: The neuroanatomical bases of envy and gloating (schadenfreude). Brain 130, 1663–1678. https://doi.org/10.1093/brain/awm093

Shamay-Tsoory, S.G., Aharon-Peretz, J., 2007. Dissociable prefrontal networks for cognitive and affective theory of mind: A lesion study. Neuropsychologia 45, 3054–3067. https://doi.org/10.1016/j.neuropsychologia.2007.05.021

Shamay-Tsoory, Harari, Szepsenwol, Levkovitz, 2009. Neuropsychological Evidence of Impaired Cognitive Emphaty in Eutyhmic Bipolar Disorder. J. Neuropsychiatry Clin. Neurosci. 21, 59–67.

Sprens, R.N., McKinnon, M.C., Mar, R. a., Levine, B., 2009. The Toronto Empathy questionnaire. J Pers Assess 91, 62–71. https://doi.org/10.1080/00223890802484381.The

Stietz, J., Jauk, E., Krach, S., Kanske, P., 2019. Dissociating Empathy From Perspective-Taking: Evidence From Intra- and Inter-Individual Differences Research. Front. Psychiatry 10, 1–8. https://doi.org/10.3389/fpsyt.2019.00126

Terrien, S., Stefaniak, N., Blondel, M., Mouras, H., Morvan, Y., Besche-Richard, C., 2014. Theory of mind and hypomanic traits in general population. Psychiatry Res. 215, 694–699. https://doi.org/10.1016/j.psychres.2013.12.042

Van Rheenen, T.E., Ganella, E.P., Bauer, I.E., Bartholomeusz, C.F., 2019. Characterization of social cognitive deficits on the schizophrenia-bipolar disorder spectrum: An overview of current evidence, Social Cognition in Psychosis. Elsevier Inc. https://doi.org/10.1016/b978-0-12-815315-4.00001-x

Van Rheenen, T.E., Lewandowski, K.E., Bauer, I.E., Kapczinski, F., Miskowiak, K., Burdick, K.E., Balanzá-Martínez, V., 2020a. Current understandings of the trajectory and emerging correlates of cognitive impairment in bipolar disorder: An overview of evidence. Bipolar Disord. 22, 13–27. https://doi.org/10.1111/bdi.12821

Van Rheenen, T.E., Meyer, D., Rossell, S.L., 2014. Pathways between neurocognition, social cognition and emotion regulation in bipolar disorder. Acta Psychiatr. Scand. 130, 397–405. https://doi.org/10.1111/acps.12295

Van Rheenen, T.E., Miskowiak, K., Karantonis, J., Furlong, L., Murray, G., Rossell, S.., 2020b. Understanding familial liability for emotion regulation difficulties in bipolar disorder. Psychol. Med. 1–8. https://doi.org/https:// doi.org/10.1017/S0033291720004626 Received:

Van Rheenen, T.E., Murray, G., Rossell, S.L., 2015. Emotion regulation in bipolar disorder: Profile and utility in predicting trait mania and depression propensity. Psychiatry Res. 225, 425–432. https://doi.org/10.1016/j.psychres.2014.12.001

Van Rheenen, T.E., Rossell, S.L., 2014. Let's face it: Facial emotion processing is impaired in bipolar disorder. J. Int. Neuropsychol. Soc. 20, 200–208. https://doi.org/10.1017/S1355617713001367

Van Rheenen, T.E., Rossell, S.L., 2013. Picture sequencing task performance indicates theory of mind deficit in bipolar disorder. J. Affect. Disord. 151, 1132–1134. https://doi.org/10.1016/j.jad.2013.07.009

Wang, Y. yu, Wang, Y., Zou, Y. min, Ni, K., Tian, X., Sun, H. wei, Lui, S.S.Y., Cheung, E.F.C., Suckling, J., Chan, R.C.K., 2018. Theory of mind impairment and its clinical correlates in patients with schizophrenia, major depressive disorder and bipolar disorder. Schizophr. Res. 197, 349–356. https://doi.org/10.1016/j.schres.2017.11.003

Watson, S., Gallagher, P., Dougall, D., Porter, R., Moncrieff, J., Ferrier, I.N., Young, A.H., 2014. Childhood trauma in bipolar disorder. Aust. N. Z. J. Psychiatry 48, 564–570. https://doi.org/10.1177/0004867413516681

Wechsler, D., 2001. Wechsler Test of Adult Reading: WTAR. Psychological Corporation, San Antonio, TX.

Young, R.C., Biggs, J.T., Ziegler, V.E., Meyer, D.A., 1978. A Rating Scale for Mania. Br. J. Psychiatry 133, 429–435. https://doi.org/10.1192/bjp.133.5.429

Table 1: Demographic and clinical variables for bipolar disorder (BD) patients and healthy controls (HC)

	HC (N=34)	BD (N=64)	Statistical test
Age, Mean (SD)[1]	35.82 (11.85)	37.11 (11.21)	$T(1, 95) = 053$, $p = .598$
Sex, N (%)			
Male	15 (44%)	33 (52%)	$Chi^2(1) = 0.49$, $p = .483$
Female	19 (56%)	31 (48%)	
Estimated premorbid IQ (WTAR), Mean (SD)[2]	114.91 (8.65)	107.69 (11.52)	U = 557.0, $p < .001$
Medication, N (%)			
Mood stabilizer	-	39 (61%)	
Antidepressants		20 (31%)	
Typical antipsychotics		2 (3%)	
Atypical antipsychotics		30 (47%)	
Age of onset, Mean (SD)[3]	-	22.48 (9.42)	
Age at diagnosis, Mean (SD)[4]	-	28.56 (10.04)	
Number of hospitalizations due to mood disturbances, Mean (SD)[5]	-	3.60 (4.99)	
Number of manic episodes (estimated lifetime), Mean (SD)[6]	-	17.62 (30.97)	
Number of depressive episodes (estimated lifetime), Mean (SD)[7]	-	25.78 (38.46)	
Bipolar type, N (%)			
Type I	-	60 (94%)	
Type II		4 (6%)	
YMRS total, Mean (SD)	-	3.39 (2.96)	
MADRS total, Mean (SD)	-	7.87 (7.20)	
History of psychosis[8]			
Yes	-	38 (60%)	
No	-	25 (40%)	

Note: T = Independent samples t-test. U = Mann Whitney U test. WTAR: Wechsler Test of Adult Reading task, YMRS: Young Mania Rating Scale, MADRS: Montgomery Asberg Depression Rating Scale. [1]Missing data from 1 bipolar disorder (BD) patient. [2]Missing data from 3 BD patients.[3]Missing data from 8 BD patients. [4]Missing data from 5 BD patients. [5]Missing data from 4 BD patients. [6]Missing data from 3 BD patients. [7]Missing data from 4 BD patients. [8]Missind data from 1 patient.

Table 2: Average Yoni performance for bipolar disorder (BD) patients and healthy controls (HC).

	HC (N=34)	BD (N=64)	Statistical test
First-order accuracy			
Cognitive	95.10 (13.32)	97.14 (10.09)	$U = 938.5$ $p = .112$
Affective	97.30 (9.56)	97.14 (8.81)	$U = 1048.5$, $p = .628$
Physical	94.85 (14.14)	91.41 (17.94)	$U = 965.5$, $p = .210$
First-order RT			
Cognitive	4226.28 (1566.72)	3438.50 (1056.26)	$U = 702.0$, $p = .004$
Affective	3667.00 (1492.20)	3293.13 (1019.07)	$U = 912.0$, $p = .189$
Physical	2528.29 (728.84)	2923.41 (1121.83)	$U = 894.0$, $p = .148$
Second-order accuracy			
Cognitive	83.33 (14.94)	75.78 (18.76)	$U = 837.5$, $p = .058$
Affective	86.52 (11.96)	79.69 (17.81)	$U = 889.5$, $p = .131$
Physical	90.20 (14.86)	88.28 (16.71)	$U = 1034.0$, $p = .651$
Second-order RT			
Cognitive	8286.55 (2863.65)	8625.06 (2668.28)	$U = 1012.0$, $p = .571$
Affective	6747.12 (1890.53)	7174.52 (2201.28)	$U = 1005.0$, $p = .536$
Physical	4796.10 (1664.39)	5317.37 (2202.29)	$U = 997.0$, $p = .497$
Fortune-of-others accuracy	82.35 (16.38)	80.21 (23.55)	$U = 1059.5$, $p = .823$
Gloat	88.24 (16.68)	76.56 (32.22)	$U = 901.5$, $p = .134$
Envy	88.73 (16.78)	90.63 (12.20)	$U = 1050.0$, $p = .750$
Identifies	87.25 (17.91)	85.68 (20.55)	$U = 1065.0$, $p = .847$
Physical #2			
Fortune-of-others RT			
Gloat	6655.23 (2935.84)	6067.81(1901.87)	$U = 1001.0$ $p = .516$
Envy	6237.19 (3051.40)	6377.92 (2617.35)	$U = 1040.0$, $p = .720$
Identifies	5294.61 (2180.84)	5051.08 (1985.81)	$U = 977.0$, $p = .407$
Physical #2	5349.72 (2239.58)	5671.00 (2264.68)	$U = 1001.0$, $p = .516$
TEQ total score[1]	49.61 (7.60)	48.13 (7.70)	T (1, 92) = -0.883, $p = .379$

Note: Values are given as mean (SD). U = Mann Whitney U test. T = Independent samples t-test. No significant group differences were observed using a Bonferroni-adjusted alpha level of $p < .002$ (.05/21) to account for multiple comparisons. The euthymic patients are included in the full bipolar disorder (BD) group. Accuracy is given in percent correct and response time (RT) is given in milliseconds. [1] Missing data from 1 BD patients and 3 HC´s.

Table 3: Inter-correlations between ToM, 'fortune of others' emotions and affective empathy in the full sample

A: Inter-correlations between accuracy variables from the Yoni task and TEQ total score

	Envy	Gloat	Identification	Physical #2	TEQ total
First-order cognitive	.174	.132	.143	-.048	.052
First-order affective	**.346**	.171	.161	.159	.126
First-order physical	.236	.131	-.021	.811	.124
Second-order cognitive	**.372**	.325	.133	.256	-.074
Second-order affective	**.463**	**.359**	.077	.235	-.062
Second-order physical	.332	.261	-.054	.251	-.077
TEQ total	-.062	-.054	.008	.086	-

B: Inter-correlations between reaction time variables from the Yoni task and the TEQ total score

	Envy	Gloat	Identification	Physical #2	TEQ total
First-order cognitive	**.371**	**.451**	**.536**	.307	-.313
First-order affective	**.411**	**.543**	**.571**	**.480**	-.193
First-order physical	.302	**.346**	.333	.262	-.161
Second-order cognitive	**.595**	**.545**	**.658**	**.638**	-.265
Second-order affective	**.439**	**.483**	**.468**	**.494**	-.224
Second-order physical	**.426**	**.456**	**.453**	**.473**	-.082
TEQ total	**-.369**	-.201	-.335	-.158	-

Note: Values represent spearman's correlations. Correlations significant at a Bonferroni-adjusted alpha level of $p < .0007$ (.05/68) are displayed in bold.