The ‘extreme female brain’: increased cognitive empathy as a dimension of psychopathology

Natalie Dinsdale, Mikael Mokkonen, Bernard Crespi

ABSTRACT
Baron-Cohen’s ‘extreme male brain’ theory postulates that autism involves exaggerated male-typical psychology, with reduced empathizing (considered here as social–emotional interest, motivation and abilities) and increased systemizing (non-social, physical-world and rule-based interest, motivation and abilities), in association with its male-biased sex ratio. The concept of an ‘extreme female brain’, involving some combination of increased empathizing and reduced systemizing, and its possible role in psychiatric conditions, has been considerably less well investigated. Female-biased sex ratios have been described in two conditions, depression and borderline personality disorder (BPD), that also show evidence of increases in aspects of empathy in some studies. We evaluated the hypothesis that BPD and depression can be conceptualized in the context of the ‘extreme female brain’ by: (1) describing previous conceptualizations of the extreme female brain model, (2) reviewing evidence of female-biased sex ratios in BPD and depression, (3) conducting meta-analyses of performance on the Reading the Mind in the Eyes test (RMET) among individuals with BPD, clinical or sub-clinical depression, and other psychiatric conditions involving altered social cognition and mood (schizophrenia, bipolar disorder, eating disorders, and autism), in relation to disorder sex ratios, and (4) evaluating previous evidence of increased empathic performance in these, and related, psychiatric conditions, and (5) synthesizing these lines of evidence into models for causes and effects of an ‘extreme female brain’. Our primary empirical results are that RMET performance is enhanced in sub-clinical depression, preserved in borderline personality disorder, and reduced in other disorders (by meta-analyses), and that across disorders, more male-biased patient sex ratios are strongly associated with worse RMET performance of patients relative to controls. Our findings, in conjunction with previous work, suggest that increased cognitive empathizing mediates risk and expression of some psychiatric conditions with evidence of female biases, especially sub-clinical depression and borderline personality disorder, in association with increased attention to social stimuli, higher levels of social and emotional sensitivity, negative emotion biases, and over-developed mentalist thought. These results link evolved human sex differences with psychiatric vulnerabilities and symptoms, and lead to specific suggestions for future work.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction
Evolved sex differences in psychological and biological traits play important roles in the development, causes, and manifestations of many psychiatric conditions (Rutter, Caspi, & Moffitt, 2003). The ‘extreme male brain’ (EMB) theory of autism postulates that autism spectrum conditions reflect extreme manifestations of ‘male-typical’ psychology due in part to high prenatal testosterone (Baron-Cohen, 2002; Baron-Cohen, Knickmeyer, & Belmonte, 2005). This theory can help to account for several notable features of autism, including its strongly male-biased sex ratio and reduced performance in measures of empathy and theory of mind (Baron-Cohen et al., 2011). Given the usefulness of the EMB theory in generating testable hypotheses and accounting for patterns in data on autism, it is of interest to investigate the other side of the spectrum: if extreme psychological ‘maleness’ can manifest in autism spectrum traits, what psychological traits and psychiatric conditions might be associated with extreme psychological ‘femaleness’?

Consideration of human psychological sex differences in the context of psychiatric conditions requires addressing two central issues at the outset. First, the ‘extreme male brain’ (EMB) and ‘extreme female brain’ (EFB) in this context have been defined psychologically and seldom involve neurological studies; the EFB and EMB, thus mainly describe psychological profiles that exist at the extreme ends of normal
distributions of sexually-dimorphic psychological traits (e.g., Grove, Bailie, Allison, Baron-Cohen, & Hoekstra, 2013). Psychological traits that differ between males and females do so statistically when measured from large samples, meaning that sex differences are small but statistically significant at the population level. Because the distributions of these psychological traits overlap considerably between the sexes, an individual male may exhibit EFB phenotypes, and likewise, an individual female may express EMB traits; however, statistically, an EMB profile is more likely to describe a male and an EFB profile is more likely in a female (Baron-Cohen et al., 2005). Average sex differences are important, however, in that they may result in the extremes of distributions of sex-differential psychological phenotypes exhibiting strong sex biases, depending on the shapes of the distributions.

Second, the extreme development of many normally-distributed traits can mediate the expression of psychiatric conditions, as for personality disorders that explicitly represent maladaptive extremes of psychological personality variation (e.g., Nettle, 2007a; Trull, 2012; Trull & Durrett, 2005; Widiger & Pernsal, 2013), and for more-severe psychiatric disorders whose psychological phenotypes grade more or less continuously into those of non-clinical populations (e.g., Constantino, 2011; van Nierop et al., 2012). For psychological and psychiatric phenotypes that are normally distributed as well as sexually dimorphic, extreme developments are thus expected to occur more often within one sex than the other and to contribute to psychological dysfunction.

In this article, we develop and evaluate central aspects of the construct of the ‘extreme female brain’ in relation to psychiatric conditions. We first briefly explain Baron-Cohen’s ‘extreme male brain’ model of autism spectrum conditions, in the context of his psychological model of low empathizing in combination with high systemizing and the male-biased sex ratios found in association with autism. Second, we describe previous conceptualizations of the ‘extreme female brain’ construct, and their relationships with empathizing, systemizing, and biased sex ratios, and our model of the EFB as proposed and evaluated here. Third, we evaluate our hypothesis through: (1) review and evaluation of evidence regarding sex ratio biases in BPD and depression, two conditions postulated as reflecting extreme female brain phenotypes that have yet to be analyzed in this context. (2) using meta-analyses to evaluate the prediction that cognitive empathizing ability, as indicated by the Reading the Mind in the Eyes test (RMET) (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001) is increased or preserved in BPD and clinical or sub-clinical depression, but not in other psychiatric conditions (bipolar disorder, schizophrenia, eating disorders, and autism); and (3) testing for a positive association between higher RMET performance in patients compared to controls, and more female-biased disorder sex ratios, across this set of disorders. Finally, we discuss the implications and limitations of our results, and make suggestions for future research.

2. Empathizing, systemizing, and the extreme male brain

A powerful framework for understanding patterns of psychological sex differences is the empathizing–systemizing (E–S) theory, which suggests that humans have evolved two parallel and complementary cognitive-affective systems (Baron-Cohen, 2002; Baron-Cohen, Richler, Bisarya, Gurunathan, & Wheelwright, 2003, 2005, 2011, Baron-Cohen & Wheelwright, 2004; Baron-Cohen et al., 2005, 2011; Chakrabarti & Baron-Cohen, 2006). By this theory, ‘empathizing’ involves the motivation and skills required to understand and interact appropriately with the social world, and ‘systemizing’ describes the drive to analyze, understand and manipulate the physical world (Baron-Cohen, 2002, 2009; Baron-Cohen et al., 2005; Lawson, Baron-Cohen, & Wheelwright, 2004; Nettle, 2007b). Crespi and Baddock (2008) and Baddock (2009) use the terms ‘mentalistic’ and ‘mechanistic’ in place of empathizing and systemizing, to capture a somewhat wider breadth of cognitive and affective systems; the terms ‘mentalizing’ and ‘motionalization’ have likewise been used extensively in clinical psychology (e.g., Choi-Kain & Gunderson, 2008; Bateman & Fonagy, 2010). Social, empathizing, and mentalistic or mentalizing cognition, and systemizing and mechanistic cognition (which are non-social), are each basically synonymous for our purposes, and these two systems are subserved by different networks of distributed regions of the brain, with activation patterns that tend to be inversely related (Jack, 2014; Jack et al., 2012). Both sets of constructs can be quantified using either self-report or task-based metrics of the relevant interacts and abilities. Given that higher fetal testosterone mediates both decreased social and empathic interest and abilities, and increased scores on metrics of systemizing (review in Baron-Cohen et al., 2011), and that their neural bases are inversely associated, empathizing and systemizing might be expected to be negatively correlated with one another, as found in some studies (e.g., Grove et al., 2013; Nettle, 2007b). However, the psychological and neurological bases of the relation of empathizing with systemizing remain to be studied in detail.

A primary sex difference in cognition, Baron-Cohen (2002, 2009) postulates, is represented by the balance of empathizing with systemizing, such that on average, males demonstrate a stronger drive to systemize, and females, on average, tend toward empathizing (Baron-Cohen, 2002; Baron-Cohen et al., 2005; Chakrabarti & Baron-Cohen, 2006; Wakabayashi et al., 2007). Under the E–S model, autism represents an extreme expression of male-typical cognition involving a strongly skewed profile of enhanced systemizing and reduced empathizing (extreme Type S in Baron-Cohen et al., 2005). This pattern of exaggerated psychological ‘maleness’ is consistent with the strong male bias in autism spectrum disorder prevalence (especially among individuals with less–severe autism), as well as with evidence linking autistic phenotypes with elevated exposure to prenatal androgens, hormones that play important roles in ‘masculinizing’ the developing brain (Baron-Cohen, 2002; Baron-Cohen et al., 2011; Lutchmaya, Baron-Cohen, & Raggatt, 2002a, 2002b; Manning, Baron-Cohen, Wheelwright, & Sanders, 2001; Wu & Shah, 2011).

Together, the E–S and EMB theories predict that a primarily-female proportion of the population will exhibit an E–S profile opposite to that observed in autism, one that is skewed toward increased empathizing and reduced systemizing (extreme Type E in Baron-Cohen et al., 2005). Consistent with this prediction, Goldenfeld, Baron-Cohen, and Wheelwright (2005) reported that, based on distributions of scores from the Empathy Quotient and Systemizing Quotient drawn from individuals with and without high-functioning autism, a small and all-female proportion of the sample (7%, and none with autism) exhibited this ‘extreme female’ profile of high empathizing and low systemizing. Similarly, data from Baron-Cohen et al. (2014) showed that ‘extreme type E’ individuals (those in the lowest 2.5th percentile for SQ and highest 2.5th percentile for EQ) demonstrated a strong female bias (with 59 of 60 individuals being female, from a total population of 2562 females and 1344 males; $\chi^2 = 28.9, P < 0.0001$; similar results are also described in Wheelwright et al. (2006). However, the question of whether or not an EFB, defined in this manner, manifests in aspects or diagnoses of psychiatric conditions has been addressed by only a small number of previous studies.

3. Previous accounts of the ‘extreme female brain’

Baron-Cohen (2002, 2012) suggested that increased empathizing drive and abilities need not negatively impact social functioning and gender-specific illness, and that low systemizing would be unlikely to cause impairment in psychological functioning. However, in these articles he did not explicitly consider extreme high levels of empathizing or combinations of high empathizing with low systemizing in the context of psychopathologies. A suite of researchers (Abu-Akel, 1999; Abu-Akel & Bailey, 2000; Crespi & Baddock, 2008; Dammann, 2003; Dinsdale & Crespi, 2013; Frith, 2004; O’Connor, Berry, Lewis, Mulherin, & Crisostomo, 2007; Sharp & Venta, 2012; Sharp et al., 2013; Zahn...
Waxler, Shircliff, & Marceau, 2008) has suggested causal roles for high levels of empathy or hyper-developed mentalistic cognition in psychotic-affective spectrum disorders, which include a set of psychiatric conditions (mainly schizophrenia, major depressive disorder, bipolar disorder and borderline personality disorder) (defined just below) that overlap substantially in their phenotypes, genetic bases, and correlates (Barnow et al., 2010; Blackwood et al., 2007; Brosnan, Ashwin, Walker, & Donaghe, 2010; Brosnan, Daggar, & Collomosse, 2010; Crespi, 2011; Glaser, Van Os, Thewissen, & Myin-Germeys, 2010; Kendler, 2005; Lieb, Zanarini, Schmahl, Linehan, & Bohus, 2004; Moritz et al., 2011; Perugi, Fornaro, & Akiskal, 2011; Slotema et al., 2012). Such overlaps, and the high comorbidities between these conditions, highlight their indistinct boundaries and partially-overlapping genetic and environmental causes. In this context, borderline personality disorder (defined below) was instead originally conceptualized as a condition at the interface of depression with schizophrenia, given its high comorbidity (coincidence in diagnosis) with depression (e.g., Luca, Luca, & Calandra, 2012), and its notable incidence (e.g., on the order of 25–50% of patients) of psychotic symptoms (Balaratnasingam & Janca, 2015).

Borderline personality disorder is a psychiatric condition that involves unstable and intense personal relationships, disturbance of self-identity, feelings of emptiness, high rejection sensitivity and expectations of abandonment, self-damaging behavior, impulsivity, anger, mood instability, and psychotic symptoms, that is highly comorbid with depression and bipolar disorder (Barnow et al., 2010; Crowell, Beauchaine, & Linehan, 2009; Fonagy & Luyten, 2005; Glaser et al., 2010; Gunderson, 1984, 2009; Leichsenring, Leibing, Kruse, New, & Leweke, 2011; Lieb et al., 2004). Dammann (2003) first, briefly, postulated a role for high levels of empathizing in this psychiatric condition, hypothesizing that borderline personality disorder in particular involves a pathological high-sensitivity to empathy, in direct contrast to the reduced empathy observed in autism. Dammann (2003, pp. 398) thus suggested that ‘one could describe mentalization disorders such as those found in BPD … as an extreme form of the female brain’. Dammann’s (2003) hypothesis was supported by a recent review analyzing performance on measures of empathy among individuals with borderline personality disorder (Dinsdale & Crespi, 2013), which reported that empathic enhancements were observed for BPD individuals (compared to controls) in 14 of 28 studies, for eight different tests of cognitive empathic skills.

Zahn-Waxler and colleagues have described a model whereby excessive levels of emotional empathy and social sensitivity, in conjunction with dysfunctional family environments and high levels of internalizing emotions (e.g., guilt, shame, sadness and embarrassment), may increase liability to depression and anxiety, primarily among adolescent and young-adult females (O’Connor, Berry, Weiss, & Gilbert, 2002, O’Connor et al., 2007; Zahn-Waxler, Crick, Shirtcliff, & Woods, 2006, Zahn-Waxler & Van Hulle, 2012; Zahn-Waxler et al., 2008). Zahn-Waxler et al. (2008) provided an account of the EMB concept in the context of this model, with a focus on early social development of girls compared to boys and potential psychological tests of high social–emotional sensitivity under stressful or abusive developmental conditions.

Crespi and Badcock’s (2008) hypothesis that psychotic-affective conditions are diametric to the autism spectrum on an axis of human social cognition, with neurotypical individuals at the center. By this model, autism is associated with under-developed social cognition, and psychotic-affective conditions (mainly schizophrenia, bipolar disorder, and major depression) with pathologically hyper-developed aspects of social cognition. Hyper-developed social cognition thus describes features of psychotic-affective conditions, such as paranoia, delusions of conspiracy, dysregulated social emotionality, and high levels of guilt and shame, in terms of hyper-mentalizing (Dziobek et al., 2006; Frith, 2004; Sharp & Venta, 2012) rather than in terms of mentalizing ‘deficits’ per se as scored by most psychological tests of social-cognitive functioning (Dinsdale, Hurd, Wakabayashi, Elliot, & Crespi, 2013). To link Crespi and Badcock’s (2008) hypothesis with Baron-Cohen’s E-S framework, the social–cognitive impairments characteristic of psychotic-affective conditions can be seen as involving (1) over-developed and dysfunctional levels of mentalistic cognition and empathy, (2) difficulties with accurate mentalizing that promote projection of extreme and incorrect mental representations onto others (Langdon, 2003; Langdon & Brock, 2008), (3) low levels of mechanistic, systematic cognition, and (4) high mentalizing in combination with other cognitive-affective alterations known to underlie pathology, such as sensory-processing deficits, altered salience, jumping to conclusions, and reduced working memory. Crespi and Badcock (2008) also summarized evidence that positive schizotypy and schizophrenia with relatively-prominent positive (compared to negative) symptoms are more common in females than males, such that hyper-mentalization may characterize female more than male cognitive profiles. Recently, a positive association has been demonstrated between ‘hyper-theory of mind’ (especially high and inaccurate mentalistic explanation) and psychotic experiences among children, with a strong female bias to hyper-theory of mind scores in one of the two samples analyzed (the one with mean participant age of 13.1, compared to 11.4) (Clemmensen et al., 2014).

Brosnan, Ashwin, et al. (2010) and Brosnan, Daggar, et al. (2010) proposed and tested the hypothesis that the EFB can be characterized by a combination of high empathizing with low systemizing, finding that among females, such an ‘empathizing bias’ was linked with increased psychotic phenotypes, especially mania and paranoia, in a non-clinical population. Larson et al. (2015) further evaluated the hypothesis and findings by Brosnan, Ashwin, et al. (2010) and Brosnan, Daggar, et al. (2010), reporting that adults with autism and psychotic symptoms showed a higher empathizing bias than adults with autism and no psychotic symptoms, with an especially strong effect for females with mania. To the extent that the ‘psychotic symptoms’ reported in this study do not represent misinterpretations (e.g., Van Slavkhlyk, Peluso, Qayyum, McPartland, & Volkmar, 2015), these studies provide evidence for a combined role of high empathizing and low systemizing in psychotic phenotypes, although the psychological mechanisms for such relationships have yet to be elucidated. In this context, Fig. 1 depicts the EMB and EFB models under the hypotheses proposed by Baron-Cohen for autism spectrum conditions, and by Crespi and Badcock (2008),

Fig. 1. By the empathizing–systemizing model extended to understanding extreme female brain phenotypes, dysfuncionally-high empathizing and low systemizing are expected to be associated with borderline personality disorder, depression, and other manifestations of psychotic-affective conditions. As such, these conditions can be considered to represent psychological ‘opposites’ to autism spectrum conditions, with regard to empathizing and systemizing.

Please cite this article as: Dinsdale, N., et al. The ‘extreme female brain’: increased cognitive empathy as a dimension of psychopathology, *Evolution and Human Behavior* (2016), http://dx.doi.org/10.1016/j.evolhumbehav.2016.02.003
Brosnan, Ashwin, et al. (2010), and Brosnan, Daggar, et al. (2010) for psychotic-affective conditions, whereby the relevant psychiatric disorders are located at opposite extremes along the empathizing–systemizing axes.

Bremer and Gallup (2012) presented a series of findings salient to the hypothesis that a combination of disordered eating with negative evaluation anxiety meets Baron-Cohen’s criteria for the EFB. Drawing on the observation that disordered eating is more prevalent among females than males, the authors surveyed a large sample of undergraduate students to investigate relationships between empathizing, systemizing, disordered eating, and attitudes toward being evaluated by others. A hyper-empathizing psychological profile predicted disordered and fear of negative evaluation by others in both sexes, and fear of being negatively evaluated was associated with lower scores in systematic thinking. Findings from objective measures of empathizing and systemizing revealed similar patterns. The EFB profile of high empathizing and low systemizing thus predicted increased risk of eating disorders and social anxiety. Further studies by the same authors found that dysfunctional attitudes toward eating, and fear of negative evaluation, were also associated with higher self-reported scores on schizotypy, including exaggerated suspiciousness, magical thinking, and paranoia. With regard to interpretation of these findings in the context of previous hypotheses, it is important to note that eating disorders are highly comorbid with depression, and with borderline (as well as schizotypal) personality disorders (Griso et al., 2003); for example, Zanarini, Frankenburg, Hennen, Reich, and Silk (2004) found that of 290 patients (77.1% female) with BPD, 86.6% also exhibited major depression, and 53.8% were diagnosed with eating disorders. Individuals diagnosed with anorexia nervosa also demonstrate increased schizotypal features (Holliday, Uber, Landau, Collier, & Treasure, 2006) and weight preoccupation is predicted by borderline personality traits (Davis, Claridge, & Cerullo, 1997). These findings indicate that eating disorders, borderline personality disorder, schizotypy, and depression, although they show notable differences, also share trans-diagnostic causes and symptoms that are associated with the psychotic-affective spectrum more generally; in this context, it is important to bear in mind that psychiatric diagnoses represent artificial constructs rather than etiologically-defined diseases per se.

Sharp et al. (2013) provided a useful definition of hyper-emotionalizing, as ‘a social-cognitive process that involves making assumptions about another person’s mental states that goes so far beyond observable data that the average observer will struggle to see how they are justified’. Sharp and Venta (2012) summarize evidence that hyper-emotionalizing is characteristic of individuals (especially females) with borderline personality disorder and borderline features, and that it may centrally involve difficulties in emotion regulation especially in social contexts. They note that despite such hyper-emotionalizing, BPD has been associated in multiple studies, predominantly of females, with superior mentalizing abilities compared to controls for some tasks (reviewed in Dinsdale & Crespi, 2013; Mitchell, Dickens, & Picchioni, 2014). Such enhanced abilities may be related to high levels of attention and vigilance in social situations (Domes, Schulze, & Herpertz, 2009; Frick et al., 2012; Sieswerda, Arntz, Mertens, & Vertommen, 2007), increased reliance on, and experience with, implicit, automatic and non-conscious social cognition (Fonagy & Luyten, 2009; Sharp & Venta, 2012; Sharp et al., 2013), and increased expression in females of psychological defense mechanisms in social contexts (Del Giudice, 2011), but such mechanisms have yet to be analyzed directly. Hyper-emotionalizing in disorders such as BPD can be consistent with increased accuracy in mentalization tasks, compared to controls, to the extent that errors are reduced on average but tend to involve over-interpretation of social stimuli (Sharp et al., 2011).

Taken together, these conceptualizations of the EFB, and sex differences in social psychopathology, all suggest that it may involve psychological dysfunctions typical of the psychotic-affective spectrum, and high aspects of empathizing or high empathizing relative to systemizing. In particular, the studies considered above suggest that BPD and depression represent among the strongest psychiatric candidates for consideration as ‘extreme female brain’ disorders, based on previous evidence regarding symptom profiles, empathic-task enhancements and female-biased sex ratios in both conditions (Dinsdale & Crespi, 2013; Mitchell et al., 2014; Zahn-Waxler et al., 2008). BPD and depression are indeed especially closely associated, with up to about 85% of individuals with BPD having comorbid diagnoses of depression (Lieb et al., 2004; Stanley & Wilson, 2006; Zimmerman & Mattia, 1999). Borderline personality features also positively predict levels of depression across non-clinical individuals (Fonseca-Pedrero et al., 2011), individuals with either BPD or depression exhibit similar five-factor model personality characteristics such as high harm avoidance and low self-directedness (Luca et al., 2012), and depression shows a strong genetic correlation with BPD (the highest among all personality disorders), indicating that these two disorders share a substantial proportion of their genetic risk factors (Reichborn-Kjennerud et al., 2010). Most notably, subclinical depression and its correlates, have also, like BPD, been associated with enhancements in empathic performance (compared to controls) across multiple recent studies (e.g., Harkness, Jacobson, Duong, & Sabbagh, 2010; Harkness, Washburn, Theriault, Lee, & Sabbagh, 2011; Poletti, Sonnoli, & Bonuccelli, 2014). These findings indicate that not only are BPD and mild or more severe depression closely associated with one another, but they also represent the only two psychiatric conditions that have been linked, in replicated studies, with enhanced empathetic expertise.

The considerations described above have motivated our primary hypothesis for conceptualization of the EFB as evaluated here: that it centrally involves empathic abilities that are enhanced, in the context of increases in socially-focused cognition and affect, but that such enhancements engender increased risk for depression and borderline personality disorder and their sub-clinical manifestations. By this hypothesis, females are thus differentially strongly affected by these two psychiatric conditions, mainly due to their increased social and empathic interests and abilities compared to males. Social and empathy-related cognitive-affective foci are thus increased to extremes in BPD and depression, with enhanced performance in some social-empathic contexts (compared to neurotypical females) but social dysfunction overall, due to hyper-emotionalizing and cognitive-affective biases. Moreover, by our hypothesis there should also be an association, among psychiatric conditions involving social cognition, between disorder sex ratios and empathic abilities, with more female-biased disorders showing relatively increased, preserved, or less-reduced empathic skills in patients compared to controls.

Our hypothesis thus predicts:

1. female-biased sex ratios in depression and borderline personality disorder;
2. evidence of empathic enhancements in borderline personality disorder and depression (clinical or sub-clinical), relative to controls (or at least preserved performance compared to other disorders); and
3. a negative correlation between empathic skills (in patients relative to controls) and disorder sex ratio (percent males), such that disorders with more female-biased sex ratios involve enhancement, or less reduction, in such abilities.

4. Sex ratios in borderline personality disorder and depression

Baron-Cohen’s extreme male brain theory for autism was inspired, in part, by the strong male biases found among individuals with autism spectrum disorders, which suggest that males are predisposed to autism as a consequence of how they differ, psychologically, from females (Baron-Cohen et al., 2011). In parallel to this reasoning, an extreme female brain theory for psychiatric conditions should, as noted above,
apply most directly to disorders that show female biases in their prevalence.

Widiger and Trull (1993) conducted a meta-analysis of demographic data from 75 studies and reported that the female to male ratio for BPD diagnoses was about 3:1. Their findings were incorporated into DSM-III (and have continued into the current DSM-V), which indicates a strongly female-biased sex ratio for this disorder (American Psychiatric Association, 2013; Sansone & Sansone, 2011; Sharp et al., 2014).

This 3:1 female bias can be considered as de facto evidence for strong female biases in BPD. This finding has, however, been subject to controversy because the meta-analysis of Widiger and Trull (1993) involved mainly clinical populations, which may be subject to biases in ascertainment or sampling (Gunderson, 2010; Widiger, 1998). Several community-based studies (which involve non-clinical populations), by contrast, provide evidence for similar rates of BPD in adult females and males (Grant et al., 2008; Jackson & Burgess, 2000). These latter studies have also, however, been criticized for methodological limitations and problems (such as use of lay interviewers) which lead, for example, to estimated overall rates of BPD much higher (about 6%) than in previous epidemiological analyses (about 0.5%–2%) (Paris, 2010).

Differences among studies in BPD sex ratios may be caused by diverse factors (Sansone & Wiederman, 2014; Skodol & Bender, 2003; Widiger, 1998). The strongest correlates of heterogeneity appear to be variation between females and males in symptom profiles, ascertainment biases due to sex differences in help-seeking, and demographic differences among the populations analyzed (e.g., Oliver, Pearson, Coe, & Gunnell, 2005; Doherty & Kartalova-O’Doherty, 2010). Most notably, males and females exhibit different sets of borderline diagnostic features (Aggen, Neale, Roysamb, Reichborn-Kjennerud, & Kendler, 2009; De Moor, Distel, Trull, & Boomsma, 2009; Distel et al., 2008; Fonseca-Pedrero et al., 2011; Furnham & Trickey, 2011; Michonski, Sharp, Steinberg, & Zanarini, 2013; Newhill, Vaughan, & DeLisi, 2010; Sansone & Wiederman, 2014; Sharp et al., 2014; Zanarini et al., 2011); for example, in one study (Fonseca-Pedrero et al., 2011), females scored significantly higher than males on measures of affective instability, fear of abandonment and negative relationships, whereas males scored higher than females for impulsiveness. Given evidence for stronger female biases among clinical than non-clinical populations for BPD diagnoses and symptoms, the borderline traits reported more commonly by females across studies (including fear of abandonment, affective instability, intense and unstable relationships, paranoia, and chronic feelings of emptiness) may engender help-seeking at lower thresholds than those traits more commonly reported by males (i.e., impulsivity and anger), and thus lead to the stronger female biases observed in most clinical settings (Oliver et al., 2005; Skodol & Bender, 2003; Widiger & Trull, 1993). Symptom profiles are especially important for inferences regarding sex ratios because to meet DSM criteria for a BPD diagnosis, an individual must exhibit any combination of five (or more) symptoms from a nine-item list, thus creating a situation where individuals with substantially different sets of symptoms (with 256 combinations in total, and overlap in as few as one criterion) may receive the same diagnostic label of BPD (Gunderson, 2010). Unbiased sex ratios in some samples assessing fit to BPD diagnoses thus appear to reflect some combination of symptom profile differences between males and females, methodological issues, and differences between the sexes in thresholds for seeking psychological help (Oliver et al., 2005; Sansone & Wiederman, 2014; Skodol & Bender, 2003). With regard to the EFB hypothesis, it is the symptom profiles that are more characteristic of females that should be associated with increased empathizing; this hypothesis has yet to be directly addressed but it is consistent with the symptom profile differences described above, and with the female biases found in all of the populations of individuals with BPD who were analyzed for RMET performance, as described below.

The evidence for strong female biases in depression is extensive and highly consistent across studies, as described extensively in previous work. Thus, two recent meta-analyses (Ferrari et al., 2013; Luppata et al., 2012), and systematic or narrative reviews (e.g., Kessler, 2003; Kuehner, 2003), demonstrate strong evidence of female biases to depression. From puberty onward, females thus experience depression more than males, and this result holds for depressive symptoms as.
well as for diagnoses of depression including dysthymia, atypical depression, seasonal depression, and rapid-cycling bipolar disorder (Diforio & Jones, 2010; Kessler, McGonagle, Swartz, Blazer, & Nelson, 1993; Lucht et al., 2003; Piccinelli & Wilkinson, 2000; Sansone & Wiederman, 2014). Overall, females are about twice as likely as males to experience a major depressive episode during their lifetime, and each year, about 12% of women compared to 7% of men receive a depressive disorder diagnosis; the higher prevalence of females diagnosed with depression each year is also mostly attributable to women having an elevated risk of first onset (Avenevoli, Swendsen, He, Burstein, & Merikangas, 2015; Kessler et al., 1993; O’Connor et al., 2007; Shibley Hyde, Mezulis, & Abramson, 2008).

The symptoms contributing to both depression and BPD exist on a spectrum and grade into normal personality variation, with both conditions sharing particular commonalities and strong positive associations with the personality trait of neuroticism (Corbitt & Widiger, 1995; Distel et al., 2008; Fossati, Borroni, Feeney, & Maffei, 2012; Sansone & Wiederman, 2014; Widiger, Trull, Clarkin, Sanderson, & Costa, 1994). Neuroticism is a normally distributed and multifaceted personality trait that most broadly indicates a person’s sensitivity to negative stimuli, and, on average, women exhibit greatly elevated neuroticism relative to men (Nettle, 2007a). This sex difference in neuroticism becomes especially pronounced in the upper tail of the distribution, where personality pathology manifests; thus, for individuals endorsing high neuroticism, on the order of 70% are female (Corbitt & Widiger, 1995; Costa & McCrae, 1988, 1992; Sansone & Wiederman, 2014).

5. Borderline personality disorder, depression, and the extreme female brain

The findings described above for female biases in depression and BPD suggest that sufficient evidence exists to merit focused investigation of the degree to which risk and expression of these psychiatric conditions may be modulated, in part, by high empathizing or some combination of high empathizing with low systemizing. Here, we focus primarily on the hypothesis that the causes of depression (clinical, sub-clinical or both) and BPD are mediated by the empathizing dimension of an EFB cognitive profile. We consider this hypothesis to be relatively productive because previous discussion of EFB models, summarized above, has focused almost exclusively on deleterious effects from increased empathizing, compared to reduced systemizing. By contrast, systemizing tasks and measures showing male biases (e.g., the Systemizing Quotient, or Intuitive Physics tests) have been investigated in few disorders other than autism, and hypothesized or empirical links between reduced systemizing and psychopathology remain almost completely undeveloped.

Of empathy measures, the Reading the Mind in the Eyes test (RMET) (Baron-Cohen et al., 2001) represents an especially useful metric for analyzing cognitive-empathic abilities across psychiatric disorders in this context. The RMET involves looking at the eye region of faces of male and female strangers, and choosing which of four words (e.g., anxious, playful, surprised, or afraid) best describes the feeling or thought portrayed; 8 can be categorized as positively-valenced, 12 as negative, and 16 as neutral (Harkness, Sabbagh, Jacobson, Chowdrey, & Chen, 2005). The test is believed to quantify some combination of rapid, intuitive matching of eye-region expressions with patterns stored in memory, with ability to explicitly identify the name of the mental state inferred (Baron-Cohen et al., 2001). The RMET is especially appropriate and useful for the analyses conducted here because females show better performance than males by meta-analysis (Kirkland, Peterson, Baker, Miller, & Pulos, 2013), the test has been used extensively among individuals with BPD, clinical and sub-clinical depression, autism, and other psychiatric conditions (much more so than any other comparable test), and enhanced RMET performance has been associated with higher scores on the Empathy Quotient (Cook & Saucier, 2010; Lawrence, Shaw, Baker, Baron-Cohen, & David, 2004; Vellante et al., 2013; Voracek & Dressler, 2006), as well as more female-typical digit ratios (van Honk et al., 2011) as an additional potential correlate of EFB psychological phenotypes. Comparable tests of abilities to decode non-verbal cues from faces also show considerable evidence of enhanced abilities among females compared to males (e.g., Hall, 1978; Hall & Matsumoto, 2004). It is important to note, however, that the RMET measures only one aspect of empathy (cognitive empathy, in the context of visually-based theory of mind abilities); ‘emotional’ empathy, which includes for example emotional resonance and empathic concern, is not addressed.

In the next section of this paper, we systematically evaluate the EFB model with regard to performance on the RMET, by analyzing RMET performance across a suite of psychiatric disorders, of which BPD and depression are predicted to best fit an EFB model. In accordance with Baron-Cohen’s conceptualization of empathizing and systemizing as mediating psychopathologies that show strong sex biases, we predict that individuals with EFB-associated psychiatric disorders should exhibit enhanced performance on the RMET or less-impaired performance than observed in other psychiatric disorders (especially autism). Moreover, under the EFB and EMB models, disorders demonstrating enhancements, or less impairment, on the RMET should exhibit relative female biases in their prevalence.

6. Reading the Mind in the Eyes in BPD, depression and other disorders

To conduct meta-analyses of RMET performance among different psychiatric conditions, we accessed Web of Science to compile a list of all publications that cited Baron-Cohen et al.’s (2001) article on the widely-used version of the RMET. Inclusion criteria comprised: (1) use of an adult version of the RMET as developed by Baron-Cohen et al. (2001); (2) application to patient and control populations, or non-clinical populations scored for psychiatric phenotypes, for all psychiatric conditions for which more than two studies were conducted; and (3) publication, or other availability, of data required to compute effect sizes. Of 1146 citing publications returned by Web of Science, 77 studies (some of which analyzed multiple independent samples, or multiple disorders) met these criteria, to produce 94 estimates: 9 for BPD, 7 for sub-clinical depression, 7 for clinical (major) depression, 10 for bipolar disorder, 14 for eating disorders (mainly anorexia and bulimia), 18 for autism, and 29 for schizophrenia. For each study (or for each independent sample within studies), we calculated effect sizes, and also recorded the sex ratios of the populations (patient populations, or all individuals for studies of non-clinical individuals). These sex ratios are not intended to represent sex ratio prevalence of the disorders investigated (although they appear to do so reasonably accurately) but they are recorded to evaluate relationships of gender to the results obtained.

Meta-analyses were conducted on the data within each of the seven psychiatric disorder categories, to quantify overall effects of the disorders with regard to RMET performance and to relate effect sizes and meta-analysis results to disorder sex ratios. Cohen’s d was calculated as the effect size by taking the mean difference between treatment and control groups, and dividing this value by the within-study pooled standard deviation (Borenstein, Hedges, Higgins, & Rothstein, 2009; Crawley, 2013). We used Cohen’s d since it would allow the inclusion of several studies that relied on effect sizes derived from correlation coefficients. Cohen’s d effect sizes were weighted by the inverse of the within-study variances. Where effect sizes were averaged for a particular study, the weights were also averaged.

Summary effect statistics were derived by dividing the sum of the weighted effect sizes by the sum of the weights. Variance of the summary effect was calculated as the inverse of the sum of the weights, and the SEM was derived by taking the square root of this summary variance. The z-statistic was calculated by taking the summary effect and dividing it by the SEM. We considered the summary effect to differ significantly from 0 (at P < 0.05) if the z-statistic was greater than 1.96. Given that the
more conservative random-effect model does not perform well with the limited sample sizes (which exist in our data for some disorders), we opted to analyze fixed-effect models. For the analyses of sex ratio, we used ANOVA to assess significant differences in sex ratio between the different disorder studies. Regression of effect size on sex ratio was performed to assess any potential relationship, whereby the standardized coefficient of the linear relationship indicated the slope of the regression line. Effect sizes used in the analyses of sex ratio were not weighted. All statistical analyses were conducted in Excel 15.0, SPSS 22.0 and R 3.1.

These analyses of RMET performance yielded three primary results. Further details are provided in Supplementary File 1 (available on the journal’s Web site at www.ehbonline.org).

First, significant negative effect sizes, indicating overall worse performance by meta-analyses on the RMET by patients compared to controls, were found for major depression, anorexia, bipolar disorder, schizophrenia, and autism (Fig. 2). Schizophrenia demonstrated the largest negative summary effect size, which was significantly lower than the effect sizes for depression (clinical and sub-clinical), bipolar disorder, and borderline personality disorder as indicated by absence of overlap for the relevant confidence intervals.

Second, BPD and sub-clinical depression were the only two conditions showing evidence of preserved (BPD) or increased (sub-clinical depression) RMET performance of patients compared to controls (non-negative effect sizes). Thus, BPD exhibited a summary effect size that was not significantly different from zero ($z = 1.24, d.f. = 8, P = 0.05$). By contrast, sub-clinical depression exhibited a significant positive summary effect size ($z = 6.25, d.f. = 6, P < 0.001$), indicating that this condition is strongly associated with enhanced performance on the RMET.

Third, the sex ratio (percent males among patients) of patients differed significantly among disorders (ANOVA: $F = 43.29, d.f. = 6, P < 0.001$), with the strongest female biases found for anorexia, BPD, and depression, and with strong male biases found for autism and schizophrenia. A linear regression of study effect size on sex ratio (percent males) was significantly negative ($F_{1,18} = 19.43, \beta = -0.425, P < 0.001$), indicating that patients, compared to controls, scored lower on the RMET in studies with relatively more male-biased patient groups (Fig. 3). Note that this effect is not due to sex differences between patient and control groups, which were almost always matched in each study.

These findings indicate that overall, more male-biased disorders tend to involve larger reductions in cognitive empathic abilities, as measured by the RMET, and that more female-biased disorders (here, BPD and sub-clinical depression in particular) tend to involve relatively spared, or enhanced, empathic abilities by this test. Given that RMET performance is also higher among neurotypical females than neurotypical males (meta-analysis in Kirkland et al., 2013), these data are consistent with the general prediction of the EFB model that such empathic abilities should be highest among females with the relevant disorder, next highest among neurotypical females, and lower still among neurotypical males; in the same way, under the EMB model, males with autism outperform neurotypical males, who outperform neurotypical females, for some tests that show well-established higher performance among males (Baron-Cohen et al., 2011). It is important to note, however, with regard to this analysis of RMET performance, that almost none of the studies presented their results separately for females and males (when both sexes were tested); more-robust tests of this EFB prediction will require analyses that take account of sex. Moreover, more male-biased disorders, especially autism and schizophrenia, are generally considered to be relatively severe (compared to the other disorders analyzed here) as regards overall cognitive deficits, which may have influenced RMET performance (although major depression, which is strongly female biased, can also be regarded as comparably severe in this general regard). Finally, we note that although anorexia exhibits strong female biases, it also involves significantly reduced performance on the RMET (Fig. 3); these findings suggest that this condition differs from borderline personality and sub-clinical depression in important ways related to theory of mind and empathic abilities, such that mentalizing may tend to be reduced rather than hyper-developed (e.g., Del Giudice, 2014; Schulte-Rüther, Mainz, Fink, Herpertz-Dahlmann, & Konrad, 2012).

How general, beyond the RMET, are findings of increases in performance on tests of empathy, compared to controls, among individuals with BPD or sub-clinical depression? Results show substantial levels of evidence for both enhanced, and reduced, performance on such tests. Overall, among 28 studies of empathic abilities in BPD reviewed by Dinsdale and Crespi (2013), about half reported evidence for enhanced mentalistic cognition in BPD, with better performance concentrated among studies that were more socially interactive. In a recent meta-analytic review of facial-emotion processing in BPD, Mitchell et al. (2014) also noted that individuals with BPD were as accurate or better than controls across a diversity of complex social judgement tasks, although there was notable heterogeneity in performance across tasks and sample populations. These findings indicate that, of all clinically-defined psychiatric and personality disorders, BPD represents the only condition to show evidence across many, diverse studies of empathic enhancements over control individuals, in a parallel way to which autism is the only disorder showing diverse evidence for enhancements over controls in visual–spatial skills and systemizing (e.g., Baron-Cohen et al., 2011; Caron, Mottron, Berthiaume, & Dawson, 2006; Mottron et al., 2013). Despite such extensive evidence of empathic enhancements in BPD, it is also important to note that a substantial proportion of studies demonstrate empathic deficits in this disorder, and that the causes of such variation in findings remain largely unclear. Moreover, although BPD did not involve enhanced RMET performance by the meta-analysis conducted here, the number of studies involved was small, which, especially given the high symptom heterogeneity of BPD, suggests that this analysis was under-powered. Additional studies of BPD are thus required for more robust evaluation of this hypothesis in relation to alternatives.

In our analysis, evidence for enhanced empathic (RMET) skills in association with depression was limited to individuals with sub-clinical manifestations and individuals with a maternal history of depression. Similarly, Lane and DePaulo (1999) found that individuals with sub-clinical depressive symptoms (dysphoria) were better than non-dysphoric individuals at detecting social deception in two experimental
paradigms, and Nettle and Liddle (2008) found that higher neuroticism predicted better performance on a test of cognitive empathy involving social story interpretation. Given that BPD, as a personality disorder, may be more similar to sub-clinical rather than major depression in the severity of its negative effects on general cognitive functioning, these findings suggest that empathic enhancements are found primarily among individuals with relatively less-severe manifestations of psychotic-affective conditions, where sex ratios also appear to be relatively female biased, as also discussed above (see also Bebbington, 1991). On the autism spectrum, sex ratios are more (male) biased toward the higher-functioning end of the spectrum, but the degree to which enhancements on tasks associated with visual-spatial skills and systemizing are differentially found among males (compared to females) with autism, or higher-functioning males with autism, remains unknown (Baron-Cohen et al., 2011, Caron, Mottron, Rainville, & Chouinard, 2004; Caron et al., 2006; Manjaly et al., 2007). Finally, the strong evidence of RMET enhancement in sub-clinical depression, but the non-significant evidence for BPD, suggests that sub-clinical depression, especially among females, may represent our best current model for EFB effects, at least in the domain of cognitive empathy and its correlates.

Despite findings that show enhancements in some aspects of empathy among individuals with sub-clinical depression, empathic skills, as indexed by the RMET, were clearly reduced among individuals with major depression. These results are concordant with findings from large sets of studies that show depression-associated deficits in cognitive and empathic functioning, in conjunction with biased attention and enhanced attention to, and recognition of, negatively-valenced emotional stimuli (e.g., Cusi, Macqueen, Spreng, & McKinnon, 2011; Inoue, Yamada, & Kanba, 2006; Leppänen, 2006; Schreiter, Pijnenborg, & Aan Het Rot, 2013). In contrast to low mood and mild depression, major depression, like many other severe psychiatric conditions, is associated with highly impaired executive function, reduced motivation and attention, and working memory deficits (Burt, Zembar, & Niederehe, 1995; Fossati, Ergis, & Allilaire, 2002; Galecki, Talarowska, Anderson, Berk, & Maes, 2015; Hasler, Drevets, Manji, & Charney, 2004). Empathizing performance in depression (as well as relatively-severe cases of BPD) is expected to be confounded by the presence of these deficits (Lee, Harkness, Sabbagh, & Jacobson, 2005); for example, Zobel et al. (2010) found that chronic depression predicted poor theory of mind performance on a cartoon-picture story task, but after controlling for logical memory and working memory, depressive status no longer predicted theory of mind ability. Studies that evaluate aspects of empathy along a continuum from very low to high levels of depression are suggested as especially useful for future work, by our results. Similar considerations regarding general deficits (e.g., Hay & Sachdev, 2011) may apply to the reductions in RMET performance found among individuals (mainly females) with eating disorders in our meta-analyses, although this hypothesis remains conjectural.

Considered together, these findings provide evidence that for the RMET and other measures of empathic abilities, higher performance by females compared to males tends to parallel higher or preserved performance by individuals with BPD or sub-clinical depression, compared to control individuals (with strong female biases among both such groups). These findings support a central prediction of the EFB model, at least with regard to its empathizing dimension. Further evaluation of the model requires additional tests, especially involving relatively realistic and socially-interactive empathy-related paradigms (Dinsdale & Crespi, 2013), individuals with sub-clinical depression or high neuroticism (e.g., Pasquier & Pedinielli, 2010) and individuals with varying degrees of severity and different symptom profiles for BPD.

### 7. Causes of enhanced or preserved mentalizing in BPD, sub-clinical depression and other conditions

In autism, a combination of enhanced perceptual, visual–spatial abilities, and systemizing, with reduced empathizing, is considered to jointly contribute to dysfunction in social interactions. By contrast, this review, and the previous work described above, have provided evidence that BPD and sub-clinical depression are associated, to some degree, with an apparently paradoxical combination of enhanced (or preserved) cognitive empathizing abilities and deficits in interpersonal social functioning, especially for BPD (Jeung & Herpertz, 2014; Lazarus, Cheavens, Festa, & Zachary Rosenthal, 2014). How can enhanced or preserved cognitive empathizing skills be reconciled with clinical or sub-clinical disorders, and its accompanying interpersonal dysfunction?

With regard to depression, several authors have described evidence that the social-cognitive features of sub-clinical depression and low mood may be interpreted as comprising a sensitive and accurate framework for focusing on and solving complex social problems, a framework that becomes disrupted in clinical and relatively-severe depression (Allen & Badcock, 2003; Andrews & Thomson, 2009; Forgas, 2007; Harkness et al., 2005, 2010, 2011; von Helversen, Wilke, Johnson, Schmid, & Klapp, 2011). By this hypothesis, increased empathy and empathic skills in sub-clinical depression may be associated with increases in attention to social cues, coupled with increased reliance on one's interpersonal social network under the conditions that generate low mood (Allen & Badcock, 2003), learning through increased practice, and rumination on social problems. A general predisposition toward the empathizing that subserves such complex social-problem-solving cognitive and affective states may be more characteristic of females than males, which would contribute to the female biases in sub-clinical and clinical depression. The apparent restriction of empathic enhancements to sub-clinical rather than clinical, depression would presumably be related to the well-established, relatively-pathological neurophysiological aspects of more-severe depression, which would obviate the expression of any forms of enhanced cognitive performance. Fig. 4a provides a hypothesized framework for conceptualizing the relationship of empathic abilities with risks and phenotypes of depression, in the context of previous studies and the ideas and evidence presented here. To the extent that sub-clinical depression indeed represents a good model for the EFB, as suggested by our RMET meta-analytic results, it would presumably involve some combination of benefits in social problem-solving with costs from deleterious aspects of depressive states.

For BPD, some early studies attributed enhanced empathy to dysfunctional parenting, whereupon children develop enhanced sensitivities to cues indicating their caregivers' mental states especially with regard to potential rejection or abuse (Carter & Rinsley, 1977; Frank & Hoffman, 1986; Krohn, 1974; Linehan, 1993; Park, Imboden, Park, Hulse, & Unger, 1992). By this hypothesis, the social difficulties characteristic of BPD result from high levels of attention and hyper-sensitivity to interpersonal interactions and reduced thresholds for emotional reactivity, which are based in part on sensitive, accurate perceptions of social cues (Dinsdale & Crespi, 2013; Miano, Fertuck, Arntz, & Stanley, 2013; Mitchell et al., 2014; Wagner & Linehan, 1999). Such abilities are underlain by some combination of intrinsically higher social and empathic sensitivities (e.g., Park et al., 1992; Frick et al., 2012), socially-challenging circumstances in childhood (including abuse) that lead to both insecure attachment and increased motivation toward socially-relevant goals, and enhanced learning of emotional cues, as demonstrated in BPD by Domes et al. (2008). Indeed, higher degrees of attachment anxiety, which are notably characteristic of individuals with BPD (Scott, Levy, & Pincus, 2009), have also been linked with increased RMET performance among healthy females (Hünefeldt et al., 2013), indicating that anxious attachment can be associated with higher cognitive-empathic performance. High sensitivity, and attention and motivation regarding social cues may also promote hyper-mentalizing (over-interpretation or imagination regarding social cues, based on one's prior expectations), which may interact with dysregulated emotionality through anxious and uncontrolled rumination, as in depression (Domsalla et al., 2013; Sharp & Venta, 2012; Sharp et al., 2011). Paranoia and extreme fears of abandonment, two facets of the diagnostic criteria for BPD, represent clear manifestations of such hyper-mentalistic
cognition. More generally, psychological accounts of the causes of BPD center on hyper-vigilance regarding the emotional states of others, negative expectations from social relationships, low thresholds for activation of social attachment systems, and low thresholds for deactivation of controlled, objective mentalization (Arntz, Bernstein, Oorschot, & Schobere, 2009; Dyck et al., 2008; Fertuck et al., 2009; Fonagy & Luyten, 2009). In the context of RMET performance, mentalization would presumably be controlled, such that empathic enhancements could be expressed in the relative absence of social–emotional dysregulation. Fig. 4b provides a conceptual model for understanding and analyzing the relationship of BPD risks and phenotypes with empathic enhancement as conceptualized by Baron-Cohen (2009). The potential roles of low systemizing under both models remain unclear.

Under the hypotheses discussed here, preserved or enhanced cognitive and empathic skills in both BPD and sub-clinical depression, thus develop primarily from biological underpinnings and social circumstances that engender increased attention and sensitivity to interpersonal cues, with social dysfunction and more-severe disorder following from some combination of increased social attention and sensitivity with emotional dysregulation, negatively-biased interpretations of social stimuli due to maladaptive developmental schema, hyper-mentalization, and increasing perceived and actual intensity of social problems (Fonagy & Luyten, 2009; Mitchell et al., 2014; O’Connor et al., 2007; Sharp et al., 2011, 2013; Unoka, Fogd, Seres, Kéri, & Csukly, 2015; Zahn-Waxler & Van Hulle, 2012; Zahn-Waxler et al., 2006, 2008). This model for BPD and depression is consistent with their high comorbidity (Lieb et al., 2004; Zimmerman & Mattia, 1999), and findings that depressive symptoms in BPD, or combined BPD and depression, are positively associated with mental state discrimination abilities in some studies, especially for negative emotions (Fertuck et al., 2009; Mitchell et al., 2014; Unoka et al., 2015). Both enhancements and dysfunction may also be related to higher intrinsic social and emotional sensitivities due to effects from differential sensitivity (Ellis, Boyce, Belsky, Bakermans-Kranenburg, & van Ijzendoorn, 2011), as postulated by Amad, Ramoz, Thomas, Jardri, and Gorwood (2014) in the context of BPD.

Potential direct and indirect roles for lower levels of systemizing (as compared to empathizing) cognition in the causes and symptoms of BPD and depression include higher reliance on automatic, emotional compared to controlled, objective non-emotional information processing during interpersonal interactions (Jeung & Herpertz, 2014; Lazarus et al., 2014; Sharp et al., 2013), and attribution of guilt, embarrassment and shame to the self in situations where such emotions are not justified by objectively-ascertained logical or moral systems (Hawes, Helyer, Herlianto, & Willing, 2013; Zahn-Waxler & Van Hulle, 2012; Zahn-Waxler et al., 2006, 2008). These possibilities remain highly conjectural, however, due to a lack of direct, focused study on aspects of systemizing in BPD or depression. These limitations notwithstanding, systemizing quotient scores are positively associated with performance on a range of visual–spatial tests (including, for example, the mental rotation test, embedded figures test, and block design) (Baron-Cohen et al., 2003; Brosnan, Ashwin, et al. (2010); Brosnan, Daggar, et al. (2010); Cook & Saucier, 2010; Carroll & Yung, 2006; Ling, Burton, Salt, & Muncer, 2009), and such tests show notable overall patterns of enhancement in autism (Auyeung et al., 2012; Baron-Cohen et al., 2011; Carroll & Yung, 2006; Falter, Plaisted, & Davis, 2008; Gilchrist et al., 2001; Soulies, Zefferio, Girard, & Mottron, 2011; Spek, Scholte, & van Berckelaer-Onnes, 2008; Stevenson & Gernsbacher, 2013) but reductions in BPD (Beblo et al., 2006; Burgess, 1990; Judd & Ruff, 1993; O’Leary, Brouwers, Gardner, & Cowdry, 1991; Stevens, Burkhardt, Hautzinger, Schwarz, & Unckel, 2004; Stone, 1992; van Reekum, 1993) and depression (Bennabi et al., 2014; Calamari, Pini, & Puleggio, 2000; Chen et al., 2013; Marcos, Salamero, Gutierrez, et al., 1994), compared to controls. With regard to robust tests of an EFB model that are fully symmetrical with the EMB model for autism (Fig. 1), BPD and sub-clinical depression should thus, in theory, be characterized by high empathizing orientation and abilities in conjunction with low systemizing (and visual–spatial) interest and abilities, which represents a strong, testable prediction.

The conceptual model for BPD and depression presented here is directly analogous to recent conceptualizations of autism, for which increases and enhancements have been observed in attention to, and perception of, non-social rather than social stimuli (Baron-Cohen, Ashwin, Ashwin, Tavassoli, & Chakrabarti, 2009; Elson, Sasson, Turner-Brown, Dichter, & Bodfish, 2012; Klin, Lin, Gorrindo, Ramsay, & Jones, 2009; Mottron & Burack, 2001; Mottron, Dawson, Soulieres, Hubert, & Burack, 2006). In autism, enhanced perceptual, systemizing, and mechanistic skills, which appear to develop in part from increased attention to non-social details in the environment as well as tendencies toward repetitive and restrictive interests and activities (Drake, Redash, Coleman, Haimson, & Winner, 2010; Happé & Vital, 2009), may contribute to disrupting the development of more complex cognitive and social–behavioral abilities, in addition to forming the basis for autistic enhancements (Baron-Cohen et al., 2009; Mottron et al., 2006; Vital, Ronald, Wallace, & Happé, 2009). The primary parallel between enhanced abilities in autism spectrum conditions, and in BPD and depression, is that in both conditions increased, selective levels of attention and interest toward particular forms of stimuli (non-social and social, respectively) may involve both the expression and the development of specialized skills and the development of deficits in central features.
and functions of cognition and affect. Autism, for example, is strongly associated with reduced attention to the eye region of faces (Tanaka & Sung, 2013); attention to the eyes is, moreover, increased in neurotypical females compared to males, and appears to be responsible in part for female superiority in recognizing facial emotions (Hall, Hutton, & Morgan, 2010). Such effects may be mediated, in part, by the mechanisms that cause inverse associations between empathizing and systemizing (e.g., Grove et al., 2013), including neurological mechanisms whereby these two systems (the ‘emotionalizing’ default network, and the task-positive networks subserving mechanistic cognition) show mutually exclusive patterns of activation (Jack, 2014; Jack et al., 2012).

Overall, interpersonal social attention, social sensitivity, interpersonal social reactivity, and engagement in social cognition more generally, appear to represent the strongest contrasts of BPD and depression with autism, and the clearest current evidence regarding the EFB hypothesis in comparison with the extreme male brain. These considerations also suggest that quantification of social-relational focus, attention, and motivation, rather than just empathizing per se, may be a more direct route to understanding the relationship between gender differences and psychopathologies linked with the EMB and EFB.

8. Conclusions, limitations and implications

We have provided here the first comprehensive theoretical and empirical framework for analysis and understanding of the EFB. Our main conclusion is that a notable body of evidence supports the hypothesis that the EFB model may be applicable to sub-clinical depression and BPD in particular, and psychotic-affective conditions more generally, just as the EMB model is applicable to autism spectrum conditions. The primary lines of evidence relevant to this inference include: (1) enhanced or preserved performance in the RMET, a paradigmatic test of cognitive empathic abilities, only among individuals with BPD or sub-clinical depression; (2) a significant correlation between disorder sex ratios and RMET effect sizes, such that female biases are associated with relatively-better performance; and (3) female biases in depression, sub-clinical depression, and in most studies of BPD. Taken together with previous work, these findings converge on indicating important roles for empathizing and mentalizing in risk and symptoms of some female-biased psychiatric conditions. More generally, these findings indicate that over-development or over-expression of adaptive, sexually-differentiated psychological phenotypes, notably visual-spatial, systemizing and mechanistic abilities in autism, and social and empathy-related abilities in borderline personality, depression, and other psychotic-affective conditions, can be associated with maladaptive cognitive traits and states. These results highlight the central importance of evolved sex differences, and psychological adaptations, in the understanding and analysis of personality variation and psychiatric conditions. In this context, further understanding of the selective, evolutionary causes of sex differences in empathy-related phenotypes would be especially useful in determining their psychological, and fitness-related, benefits and costs.

The primary limitation of the analytic aspect of our study is its restriction to one test of cognitive empathy, the RMET, which is due to this being the only such test that has been performed across enough disorders and conditions for meaningful, synthetic tests to be performed. Moreover, numerous of studies are small for some disorders (including BPD, sub-clinical depression, and depression), which reduces statistical power and increases effects from possible heterogeneity among studies in methodology. Future studies should also consider the roles of increased affective empathy in female-biased psychopathology (Zahn-Waxler et al., 2006, 2008), especially in the context of how higher social sensitivity among females (in contrast to reduced social sensitivity in autism, and in males) may contribute to risks of depression and BPD.

Our analysis of theory and evidence regarding the EFB has several implications for clinical work, and for specific future studies that would provide further tests of predictions regarding EFB models and conceptualizations. First, Sharp et al. (2013) found, with regard to borderline personality traits in adolescents, that hyper-mentalizing, but not other measures of social-cognitive reasoning, exhibited malleability through mentalization-based therapies. This finding suggests that dysfunctionally-increased empathy-related psychological phenotypes may provide more-effective targets for therapy than social–emotional or social-cognitive ‘deficits’ that are usually considered to involve reductions of trait expression. Second, given that aspects of female gender ‘protect’ females from autistic impairments (Robinson, Lichtenstein, Ankarsäter, Happé, & Ronald, 2013), some aspects of male gender may also tend to ‘protect’ males from depression and BPD; determining the nature of such protective factors may provide clues for improving preventative and therapies. McHenry, Carrier, Hull, and Kabbaj (2014) suggested, for example, that testosterone may have protective benefits against anxiety and depression. In this context, our hypothesis also predicts that depression and BPD should be associated with low prenatal testosterone (or low prenatal testosterone relative to prenatal estrogen), in direct contrast to the relatively-high prenatal testosterone associated to some degree with autism (Baron-Cohen et al., 2003; Lutchmaya et al., 2002a, 2002b; Lutchmaya, Baron-Cohen, Raggatt, Knickmeyer, & Manning, 2004). Third, major gaps remain in the literature on the roles of empathizing and systemizing in BPD, depression and other psychotic-affective conditions, especially with regard to (1) how these psychological dimensions may be related to dysfunction through effects of high empathizing combined with low systemizing, (2) how clinical and sub-clinical psychotic phenotypes are related to empathic interests and abilities in females and males (Brosnan, Ashwin, et al. (2010); Brosnan, Daggar, et al. (2010)), and (3) how females diagnosed with BPD or depression differ from males diagnosed with BPD or depression with regard to empathic skills. Given the success of the EMB model in advancing our understanding of autism, increased study of hypotheses motivated by its logical opposite, an EFB framework, should generate novel perspectives and insights for addressing these and other questions.

Acknowledgments

We are grateful to K. Harkness, M.-C. Lai, K. Levy, M. Lombardo, M. Sabbagh, S. Unoka and L. Scott for providing us with raw data for RMET effect size calculations, to Silven Read for technical help, to M. del Giudice and three anonymous reviewers for helpful comments, and to NSERC and the Academy of Finland for financial support.

References


