
Sharma AN, Barron E, Le Couteur J, Close A, Rushton S, Grunze H, Kelly T,
Ferrier N, Le Couteur AS.

[Facial emotion labeling in unaffected offspring of adults with Bipolar I
Disorder.](#)

Journal of Affective Disorders 2017, 208, 198-204.

Copyright:

© 2016. This manuscript version is made available under the [CC-BY-NC-ND 4.0 license](#)

DOI link to article:

<http://dx.doi.org/10.1016/j.jad.2016.10.006>

Date deposited:

09/12/2016

Embargo release date:

21 October 2017



This work is licensed under a
[Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International licence](#)

Facial emotion labeling in unaffected offspring of adults with Bipolar I Disorder

Aditya Sharma PhD

Evelyn Barron MRes

James Le Couteur MRes

Andrew Close PhD

Steven Rushton PhD

Heinz Grunze PhD

Thomas Kelly PhD

Nicol Ferrier MD (Hons.)

Ann Le Couteur FRCPsych

Abstract

Background

Young people 'at risk' for developing Bipolar Disorder have been shown to have deficits in facial emotion labeling across emotions with some studies reporting deficits for one or more particular emotions. However, these have included a heterogeneous group of young people (siblings of adolescents and offspring of adults with bipolar disorder), who have themselves diagnosed psychopathology (mood disorders and neurodevelopmental disorders including ADHD).

Methods

24 offspring of adults with bipolar I disorder and 34 offspring of healthy controls were administered the Diagnostic Analysis of Non Verbal Accuracy 2 (DANVA 2) to investigate the ability of participants to correctly label 4 emotions: happy, sad, fear and anger using both child and adult faces as stimuli at low and high intensity.

Results

Mixed effects modelling revealed that the offspring of adults with bipolar I disorder made more errors in both the overall recognition of facial emotions and the specific recognition of fear compared with the offspring of healthy controls. Further more errors were made by offspring that were male, younger in age and also in recognition of emotions using 'child' stimuli.

Limitations

The sample size, lack of blinding of the study team and the absence of any stimuli that assess subjects' response to a neutral emotional stimulus are limitations of the study.

Conclusions

Offspring (with no history of current or past psychopathology or psychotropic medication) of adults with bipolar I disorder displayed facial emotion labeling deficits (particularly fear) suggesting facial emotion labeling may be an endophenotype for bipolar disorder.

Text

Introduction

Emotional processing deficits (as assessed using tasks designed to test the ability to label facial emotions) have been described in young people with Bipolar Disorder (BD) both during affective episodes and when euthymic (Dickstein et al., 2009). These findings have led to the suggestion that disturbances in emotional processing capacity may be a factor in the development of BD (Yurgelun-Todd et al., 2000). It has also been proposed that difficulties in regulation of affect and interpretation of emotion could co-exist in BD. The precise mechanism is unclear with researchers suggesting that these difficulties may lead to BD, be a consequence of BD or that both problems operate through shared specific brain areas (George et al., 1998). A recent meta-analysis compared the voxelwise convergence of fMRI findings when analysing facial emotion recognition in young people with BD, adults with BD in comparison with healthy controls using Activation Likelihood Estimation (ALE) (Wegbreit et al., 2014). Based on their report of significantly greater convergence of amygdala hyperactivation among young people with BD when compared with adults with BD, the authors suggested that hyperactivation of systems involving the amygdala, prefrontal, and visual areas may be important in the emotional dysfunction present in young people with BD. However, it is not just patients (child and adult) with BD who have been shown to have these deficits as individuals at risk for BD (including Offspring of Bipolar Parents (OBP)) have also been shown to have deficits in identifying and labeling facial emotions correctly (Brotman et al., 2008a; Brotman et al., 2008b; Hanford et al., 2016; Whitney et al., 2013). This leads to the intriguing possibility that perhaps these deficits in the ability to correctly label facial emotions seen in young people at risk for BD may be a potential endophenotype for BD (Brotman et al., 2008a; Brotman et al., 2008b; Hanford et al., 2016; Whitney et al., 2013).

Brotman and colleagues used the Diagnostic Analysis of Non Verbal Accuracy 2 (DANVA 2) (Nowicki and Carton, 1993) to test the hypothesis whether young people 'at risk' for BD (with a family history of BD) have facial emotion labeling deficits similar to those seen in subjects with Paediatric Bipolar Disorder (PBD) (Brotman et al., 2008a). The 'at risk' subjects included siblings of individuals with Narrow Phenotype Bipolar Disorder (NPBD) (Leibenluft et al., 2003) (Bipolar I Disorder (BDI) and Bipolar II Disorder (BDII)) (n=14) and OBP (BDI and BDII) (n=10). In their study, any 'at risk' subjects with a current or past history of mood disorders were excluded from the analysis; however, other co-morbidities such as Attention Deficit Hyperactivity Disorder (ADHD) were not excluded. IQ was measured using the Wechsler Abbreviated Scale of Intelligence 4 (WASI 4) (Wechsler, 1999). Whilst subjects in both NPBD and 'at risk' group had high rates of psychopathology, only subjects in the NPBD group were on medication. The authors conducted primary analyses using ANOVA to assess group differences in age and IQ. As the data included observations that were non-independent (with more than one child per family) the authors used a linear mixed modelling approach as a secondary statistical technique. They reported that subjects in both the NPBD and the 'at risk' for BD groups made statistically significantly more errors labeling emotions in both child and adult face stimuli than subjects in the HC group. A sub analysis in the 'at risk' group, no differences in the number of errors on DANVA 2 between OBP and siblings of NPBD were found. Further, an analysis of at risk youth without ADHD or anxiety revealed that their findings were not influenced by the present of psychopathology in these participants

Brotman and colleagues subsequently assessed emotion processing using the Emotional Expression Multimorph Task (EEMT) (Blair et al., 2001) in young people with NPBD, 'at risk' for BD and HC group to assess ability to label facial emotions (Brotman et al., 2008b). The EEMT included the morphing of the facial stimuli (valid and reliable Ekman pictures of adults showing a variety of facial expressions (Ekman and Friesen, 1976)) from neutral (0% intensity) to the prototypical emotional expression (100% intensity). The subject being assessed could stop the morph at any stage allowing the authors to detect the intensity needed to correctly identify facial emotion instead of relying on the 'high-low' intensity paradigm used in other assessments such as DANVA 2 (Nowicki and Carton, 1993). The disadvantage of the EEMT is that it does not have any child faces as stimuli. As in the previous study the subjects in the 'at risk' for BD group had high rates (28%) of psychopathology; however, none of these subjects were on any psychotropic medication. The authors reported that compared to subjects in the HC group, subjects in both NPBD group and 'at-risk' group, required higher emotional intensity before correctly identifying the emotion being displayed. The performance of subjects in the 'at-risk' for BD group and NPBD group did not differ. For both analyses, the group-by-emotion interaction was not significant, indicating that face emotion type did not moderate group differences. This would imply that there was no specificity for subjects to make errors in labeling specific facial emotions on the EEMT.

In a study, by Whitney and colleagues, differences in socio-emotional processing and functioning in OBP and Offspring of Healthy Controls (OHC) (Whitney et al., 2013) were investigated. The authors recruited participants in the OBP group who all had psychopathology (mood disorders, ADHD). The mood disorder symptoms were reliably identified by using the Washington University in St. Louis Kiddie Schedule for Affective Disorders and Schizophrenia (WASH U KSADS) (Geller et al., 1996) and the Kiddie Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime version (KSADS-PL)(Kaufman et al., 1997) was used to identify non-mood disorders symptoms. The OHC group were defined as having no personal/family history of mental health disorder. The IQ of the offspring in both groups was assessed using the WASI (Wechsler, 1999). The other measures included DANVA 2 (Nowicki and Carton, 1993) and the Affect Recognition (AR) subtest of NEPSY II (Korkman et al., 2007). The authors reported that the OBP and OHC groups showed no statistically significant differences on either the scores of AR (NEPSY II) nor the emotion labeling using both child and adult faces on DANVA 2. Further, the rates of errors in OBP did not differ from previous studies by Brotman and colleagues (Brotman et al., 2008a; Brotman et al., 2008b). The rates of errors in OHC group in their study, however, were higher than in the previously reported studies (Brotman et al., 2008a; Brotman et al., 2008b) which may go some way to explain the lack of significant difference between the OBP and OHC groups.

In the most recent study by Hanford and colleagues DANVA 2 (Nowicki and Carton, 1993) was utilised to compare emotion labeling in OBP (n:18 had diagnosed psychopathology, age: 13.8±2.6 years, 44% female and n:12 were free from psychopathology, age: 12.8±3.0 years, 42% female) and age- and sex-matched OHC (n:20, age: 13.3±2.5 years, 45% female) (Hanford et al., 2016). The authors utilised general linear models to report that compared to OHC, both OBP (diagnosed with and free of psychopathology) groups made more errors on the adult face task ($p_{\text{cor}}=0.014$). The OBP group were 2.3 times [90%CI:0.9–6.3] more likely and 4.3 times [90%CI:1.3–14.3] more likely to make errors on sad and angry faces, respectively. The study was well designed with appropriate statistical techniques but it was not clear whether the OBP group included probands of BDI and/or BDII.

Although these studies have advanced the understanding of emotional processing in youth at risk for BD, all have methodological limitations including: small heterogeneous sample sizes of 'at risk' individuals, a range of BD diagnoses in the bipolar probands (adolescents with NPBD, adults with BD) which included both BDI and BDII, inclusion of subjects with existing psychopathology and use of psychotropic medication to manage the psychopathology. It is unclear whether the lifetime rates of illness in subjects 'at risk' for BD varies according to the type of relative (offspring or sibling) (Craddock and Jones, 1999). Gershon and colleagues initially reported that the risk in siblings could be greater than that in offspring (Gershon et al., 1975), however, in a later study the risk was reported as greater in offspring (Gershon et al., 1982). Perhaps, because of small sample sizes or because there appears to be no consistent difference between risks in offspring and siblings, published studies to date have reported only the pooled estimate of risk for all types of at risk subjects (Craddock and Jones, 1999). This study aims to advance understanding of facial emotion labeling in bipolar disorder by addressing previous research limitations. The hypothesis to be tested was that OBP will demonstrate more errors on the facial emotion labeling task particularly for low intensity stimuli and stimuli with child faces and also investigate for any contributions made by specific sociodemographic variables to the variance of errors in facial emotion labeling. Further, the OBP would only include a younger age range (6-14 years) who had not developed any mental health disorder or co-existing neurodevelopmental disorder and whose parents had BDI (thereby reducing heterogeneity of the bipolar proband).

Methods

This study took place in the North East of England, UK. Ethical approval was obtained from Northumberland Research Ethics Committee (ref 08/40902/12).

Participants

OBP: Consultant Psychiatrists in Community Adult Mental Health Teams approached adult patients with BDI who had their biological offspring (age 6-14 years) living with them. Written information sheets (adult version for parent with bipolar disorder and child version for their offspring) and an Expression of Interest (EOI) form were given to these patients. OHC were recruited through primary and secondary schools in the North East of England with the aim of recruiting a sample matched on age, gender and socioeconomic status (SES). Appropriate versions of the written information sheets (adult version and child version) and an EOI Form were provided. Once the research team received an EOI, written informed consent was obtained from the parent and assent from any child 10 years or older recruited to the study. The main inclusion criterion for subjects in the OBP group was having one biological parent with a diagnosis of BDI. The main inclusion criterion for subjects in the OHC group was having biological parents and first degree relatives with no history of mental health disorders. In addition, subjects in both OBP and OHC groups were required to meet the following common inclusion criteria: age range: 6 to 14 years of age at the time of testing, FSIQ \geq 70 and sufficiently familiar with the use of the English language to allow them to undertake the facial emotion labeling task. The exclusion criteria for subjects in the OBP and OHC groups included the presence of a currently recognised medical condition and/or substance use/dependence that would impact on the facial emotion labeling task.

Assessments

Parental diagnosis of BDI was confirmed using the Structured Clinical Interview for DSM-IV Disorders (SCID) (First et al., 2002). Control families were interviewed to exclude any personal history or family history within first degree relatives of psychiatric disorders. Parental Socio-Economic Status (SES) was ascertained using the Hollingshead and Redlich Scale (Hollingshead and Redlich, 2007). An assessment of psychopathology in the offspring was completed with the parent using the WASH-U-KSADS (Geller et al., 1996). IQ was assessed using the Wechsler Abbreviated Scale of Intelligence (WASI) (Wechsler, 1999). Facial emotion labeling was assessed using the Diagnostic Analysis of Nonverbal Behavior (DANVA 2) (Nowicki and Carton, 1993). This measure has been used in studies with children (including children in the preschool age range) (Nowicki and Mitchell, 1998). It includes both child and adult faces at high and low intensity as stimuli. It assesses the ability to correctly label 4 emotions: happy, sad, fear and anger. The two subtests in the DANVA 2 were: Adult Facial Expressions Test (DANVA 2-AF) and the DANVA 2--Child Facial Expression Test (DANVA 2-CF). The DANVA 2-AF includes 24 photographs of adult facial expressions of emotions and similarly the DANVA2-CF consists of 24 photographs of child facial expressions of emotions. The facial expressions stimuli include an equal number of male and female faces and high- and low-intensity faces. For the present study, each stimuli were presented to participants on a laptop computer and after 2 seconds the 4 possible emotions came on the same screen as options. The participants had to choose the emotion option they thought was correct. All participants had the stimuli presented in the same order. Responses were recorded on the laptop computer. The possible range for scores on accuracy for the DANVA2-AF is 0 to 24 and for the DANVA2-CF is 0 to 24. Scores on the DANVA 2 have not been found to be related to IQ scores or tests of general cognitive ability in young people (Baum et al., 1996). However, given the lack of any published data in the field, a decision was made to use IQ scores as a covariate in the statistical analysis in the facial emotion labeling scores. Further, the few other studies that have been published findings for subjects 'at risk' for BD, have included IQ as a covariate in the statistical plan of analysis to compare scores of facial emotion labeling between OBP and OHC (Brotman et al., 2008a; Brotman et al., 2008b).

Procedure

The assessments were all conducted on the same day at similar times in the day and in the same order for all participants to minimise variability. All measures were administered by the research team trained to research reliability. Fortnightly supervision meetings were used to ensure reliability during the study period.

Data analysis

Mixed-effects models were used to assess differences in the facial emotion labeling of OBP and OHC. To take into account the multiplex (more than 1 offspring per family) nature of recruitment of both OBP and OHC, families were grouped according to the bipolar status of the parents. IQ, age (in months) and gender of the offspring as well as SES of the family were included as independent explanatory variables (fixed-effects). In addition, the following variables were included as categorical factors in order to explain variation in facial emotion labeling: bipolar status of parent, emotional intensity of stimuli (high vs low) and whether the stimuli used children or adult faces. In order to encapsulate the hierarchical nature of the study design, a specified nested random-effects structure (child nested within each family) was utilised to allow for unexplained variation to be partitioned to an appropriate scale. Furthermore, a correlation structure was incorporated into models that took into consideration that observations within each hierarchical, or grouping level, may be correlated. All analyses were undertaken in the 'R' package for Statistical computing, release R 3.3.1 (R Core Team,

2016). To generate the mixed-effects models, the Linear and Nonlinear Mixed-effects Models package was used (Pinheiro et al., 2016)

Results

Recruitment, Clinical and Sociodemographic variables

Thirty-four OBP from 25 families where one parent had BDI were identified. Of the families identified, 2 families did not complete the EOI form, 1 family did not consent and 2 families withdrew consent during the study. Five (17%) OBP (1 had Autism Spectrum Disorder (ASD), 1 had ASD and ADHD, 1 had BD and ASD 1 had ADHD and 1 had Post Traumatic Stress Disorder) from 3 families who displayed psychopathology on the assessment using the WASH-U-KSADS were excluded. None of the OHC recruited and assessed, displayed any psychopathology. The final dataset consisted of 24 OBP and 34 OHC. The sociodemographic characteristics are presented in table 1. Groups were matched on age and there were no significant differences between the groups on gender, SES (Hollingshead and Redlich Scale) and total number of children in the family.

Insert table 1 about here

DANVA 2: Assessment of Errors in Emotion Labeling in OBP vs OHC

The number of errors made in labeling facial emotions between the OBP and OHC groups and the sociodemographic variables under study are shown in table 2. The number of total errors in the OBP group (mean: 12.13; SD 7.59; range: 3-33) was statistically significantly higher ($p=0.05$) than the number of errors in the OHC group (mean: 10.03; SD 4.82; range: 4-26). SES of the family and IQ did not significantly affect the average numbers of errors recorded even though OBP and OHC groups differed statistically on IQ scores ($p=0.001$). Age of the subject contributed to the model ($p=0.004$) implying that the older the child the lower the number of errors made in recognising facial emotions. There were also significantly higher errors associated with the presentation of child faces in contrast to adult faces ($p=0.0004$). The intensity of stimulus also significantly increased the average number of errors recorded ($p<0.0001$) indicating that children made a higher number of errors when shown stimuli of low intensity as when compared to stimuli of high intensity. In addition, the type of emotional stimuli was controlled for and comparisons made between the average number of errors associated with fearful, happy and sad emotions to those associated with angry emotions. In all cases, fearful, sad and happy emotions were associated with a significant reduction ($p<0.0001$) in the average numbers of errors compared to the errors associated with angry emotions.

Insert table 2 & 3 about here

DANVA 2: Errors by emotion in OBP vs OHC

The number of errors whilst labelling fearful faces in OBP group (mean: 4; SD 3.2; range: 1-12) was significantly higher ($p=0.04$) than in OHC group (mean: 3.09; SD 2.4; range: 0-11). The number of errors whilst labelling angry faces in OBP group (mean: 4.41; SD 2.11; range: 1-9) was higher than in OHC group (mean: 3.91; SD 1.93; range: 0-9); however, this was not significant ($p=0.21$). The number of errors whilst labelling sad faces in OBP group (mean: 2.04; SD 3.2; range: 0-10) was not statistically different ($p=0.64$) from OHC group (mean: 2.12; SD 2.8; range: 0-8). The number of errors whilst labelling happy faces in OBP group (mean:

1.21; SD 1.97; range: 0-8) was higher than in OHC group (mean: 0.79; SD 1.31; range: 0-4) but this did not attain statistical significance ($p=0.15$). Furthermore, subjects made more errors when presented with stimuli of sad and fearful child faces. The subjects made more errors in identifying angry, fearful and happy emotions of low intensity. These findings are represented in Figure 1. Further analyses were conducted to assess the effect size of the contribution made by parental BD to errors in labeling facial emotion. A post-hoc analysis was done using the variance inflation factors to account for the multivariate design and multi-level modelling (Hsieh et al., 2003) to calculate the actual power obtained by each statistical test. Using this approach, the effect size for all 4 emotions combined was 0.525 with a power of 0.86. For angry faces the effect size was 1.139 with a power of 0.999 and for fearful faces the effect size was 0.844 with a power of 0.8. On the other hand, for happy faces the effect size was 0.112 with a power of 0.185 and for sad faces the effect size was 0.366 with a power of 0.475.

Discussion

Key finding

This study reported that bipolar status of parent had a statistically significant contribution to the number of errors made by the offspring assessed in this study i.e. OBP made more errors in labeling and identifying facial emotion (across all 4 emotions *viz.* happy, sad, fearful and angry) with a medium effect size (Cohen's $d=0.34$). This finding is a replication of the previously reported studies (Brotman et al., 2008a; Hanford et al., 2016). Whitney and colleagues, however, did not report this finding (Whitney et al., 2013). However the high error rates in the OHC group in their study (Whitney et al., 2013) could help explain the lack of a statistically significant difference. This study has also added new evidence by increasing the specificity of the previous findings as the 'at risk' group in this study included only OBP in contrast to the study by Brotman and colleagues (Brotman et al., 2008b) which included both OBP ($n=10$) and siblings of adolescents with BDI ($n=14$) and the study by Hanford and colleagues which included OBP (both with and without psychopathology) (Hanford et al., 2016). As stated before, remains unclear whether the lifetime rates of illness in subjects at risk for BD varies according to the type of relative (offspring or sibling) (Craddock and Jones, 1999). However, at present there is persisting uncertainty about whether or not the degree of genetic loading and vulnerability in siblings of adolescents with BDI and OBP is different. For this reason, reporting research findings separately for the 2 groups is still important.

Implications of errors in labeling fear in OBP

Further analyses, identified that OBP made more errors labeling fearful facial emotions compared with OHC ($p=0.04$; Cohen's $d=0.34$). This finding has previously been reported in studies of adults with BD (Lembke and Ketter, 2002; Venn et al., 2004) and in adolescents with BD (Rich et al., 2006). In a recent study by Olsavsky and colleagues, subjects with PBD, subjects 'at risk' for BD (siblings of subjects with PBD and OBP) and HC were asked to rate their fear of fearful faces (Olsavsky et al., 2012). In their fMRI study both PBD and unaffected at-risk subjects exhibited amygdala hyperactivity as compared to HC. The authors hypothesised that amygdala activation specifically to fearful faces may be an endophenotype for BD. However, in the two

other studies by Brotman and colleagues of youth 'at risk' for BD, errors when labeling specific facial emotions were not reported (Brotman et al., 2008a; Brotman et al., 2008b). Roberts and colleagues studying youth (18- to 30-year-olds) at risk for BD demonstrated reduced brain signal of the left Inferior Frontal Gyrus (IFG) when inhibiting responses to fearful face stimuli, compared with subjects from control families (Roberts et al., 2012). This is an interesting finding when considered alongside the data from this study. One interpretation might be that the failure to inhibit responses to fearful faces in the study by Roberts and colleagues could be as a consequence of an inability to recognise fear (Roberts et al., 2012). This might in turn add support to the hypothesis that the aetiology for the development of BD may lie in the neural (dys)connectivity between prefrontal and subcortical limbic structures. This proposed deficit in the neural disconnect could be through a lack of inhibition to the limbic structures (in particular the amygdala largely implicated in recognition of fear).

Implications of errors labeling low intensity stimuli

This project reported more facial emotion labeling deficits when subjects were presented with low intensity stimuli ($p < 0.00001$). These findings are similar to previous reports by Brotman and colleagues (Brotman et al., 2008a; Brotman et al., 2008b). This is an important finding as in real life situations, people often show subtle variations in expression; whilst rarely displaying extremes of emotion (Schepman et al., 2012). OBP, however, are exposed to extremes of emotions when their parents are in affective episode. The higher rates of errors on 'low intensity' stimuli in the OBP sample in this study may indicate an inability to identify the subtlety in emotional variations of real life. This, in turn, could impact on psychosocial function in such a sample. A study of adult patients with BD who had residual symptomatology reported impaired social cognition including a social cognitive bias toward hostility, intention, anger, and aggressiveness and increased facial emotion labeling errors compared to a healthy controls (Lahera et al., 2015). The authors postulated that this impaired social cognition may contribute to psychosocial dysfunction associated with BD. Other researchers have reported that facial emotion recognition may protect quality of life in BD (Fulford et al., 2014). It has been suggested that cognitive remediation could be an important developmentally appropriate and important treatment to improve emotion recognition abilities in young people with BD (Wegbreit et al., 2015).

Implications of errors labeling affect using child faces as stimuli

More errors were made by OBP than OHC when subjects were presented with 'child faces' stimuli ($p = 0.0005$). This has been previously reported in studies assessing adolescents 'at risk' for BD (Brotman et al., 2008a) as well as adolescents with BD (McClure et al., 2003). This is an important finding that should influence choice of test stimuli in further research studying facial emotion labeling in adolescents at risk for BD as the absence of findings in tasks employing only adult faces as stimuli may be due to this finding. Further, research that attempts to assess the utility of cognitive remediation strategies such as the one being proposed by the NIH group: NIMH Child Emotional Faces Picture Set (NIMH-ChEFS) (Egger et al., 2011) should employ age and developmentally appropriate stimuli. Young people spend considerable amounts of time interacting with peers. Deficits in the ability to identify facial emotions could add to the vulnerability in OBP (as in this study) and possibly impact on the trajectory of emotional processing in OBP.

IQ and face emotion labeling

IQ did not help explain the variance in the sample under study in this project. This is perhaps not surprising as the authors of the DANVA 2 report that errors made in facial emotion labeling are not influenced by IQ. However, given the lack of published data at the time of the start of this study, an *a priori* decision had been made to include IQ in the analysis. Other studies using the DANVA 2 in 'at risk' subjects have used IQ as a covariate in analyses (Brotman et al., 2008a). The study by Baum and colleagues however did not report any contribution made by IQ scores to the ability of individual to label facial emotion (Baum et al., 1996).

Age and gender of participants

Age and gender (male) of the subjects made a statistically significant contribution to the number of errors. Some studies have identified that the ability to recognise and label facial emotions continually improves with increasing age (Boyatzis et al., 1993). Further, data from typically developing subjects indicates that the ability to recognise *different facial emotions* develops at different stages in the lifespan (De Sonneville et al., 2002). However, other authors have reported that face emotion labeling deficits are not related to the age of onset in adults with BD (Bozikas et al., 2006) or with current age at time of assessment in NPBD (Rich et al., 2008). Some authors have stated that age (particularly older than 6 years) does not improve face emotion labeling (for the 6 basic emotions) (Markham and Adams, 1992; McClure, 2000). There is, therefore, uncertainty about the impact of age on face emotion labeling. Turning to gender and its association with face emotion labeling, a meta-analytic review by McClure focussing on facial emotion processing in infants, children and adolescents reported that males made more errors in labeling facial emotion (McClure, 2000). In the study by Rich and colleagues, using the EEMT, male subjects with NPBD too correctly identified facial emotions at a higher intensity than female subjects (Rich et al., 2008). Further studies investigating facial emotion labeling should bear these factors (age and gender of subjects) in mind, when choosing facial emotion labeling tasks and designing study protocols. Furthermore, subjects made more errors when presented with child faces as stimuli when the emotion to be recognised was sad and fearful. Subjects of a male gender making more errors labeling affect has been previously described in young adults as well as adolescents (McClure, 2000; Mufson and Nowicki, 1991).

Strengths

First the parents with BDI (and their offspring) were identified from community clinics rather than specialist tertiary healthcare services. The findings from the study may therefore be more likely to be representative of and relevant for the wider population of individuals with BD particularly in the UK as previous studies are from the USA. Second steps were taken with the aim of reducing both behavioural and genetic heterogeneity and increase the likelihood of studying a relatively homogenous sample of both OBP and OHC groups. Only adults with BDI were approached to take part in the study. The current DSM 5 classification criteria differentiate BDI and BDII as distinct disorders (American Psychiatric Association, 2013). Research suggests that these disorders may differ in the severity of neurocognitive deficits, social cognitive deficits and psychosocial impairment although further research is needed as the precise details remain uncertain (Judd et al., 2005; Simonsen et al., 2008; Torrent et al., 2006; Wingo et al., 2010). Thus studying samples of individuals at risk

for both BDI or BDII may inadvertently introduce clinical/biological heterogeneity. Next only offspring of parents with BDI were included, rather than using a combined sample of siblings of adolescents with BD and offspring of parents with BD. Finally, subjects with any current and/or past history of mental health disorders were not included in analysis as the diagnosed psychopathology could contribute to the findings. These attempts increase the possibility that any identified profile of facial emotion labeling deficits in the OBP may inform subsequent investigation of underlying psychopathology and search for a potential endophenotype.

Limitations

These include the small sample size which was however, comparable to previously published work in this field. Further, 80% statistical power was achieved for all errors combined and average errors on labeling fearful and angry faces but not for the analyses of the errors on labeling happy and sad faces. The next step forward would be for larger collaborative multi-site studies. A second limitation was that the study team were not blind to the group that the subjects belonged to. This lack of blinding could have introduced a bias and as a result influenced the findings. However, the use of objective measures in the assessment battery makes this unlikely. Thirdly, a limitation of the DANVA 2 is the absence of any stimuli that assess subjects' response to a neutral emotional stimulus. Ideally the study should have had a measure of the OBP ability to label neutral emotions but at the time of the study development there were no tasks that assessed a wide range of facial emotions using both child and adult faces, low and high intensity that also incorporated neutral emotions. Although OBP with recognised current/lifetime psychopathology were excluded from the analyses, the project did not have rating scales that assessed for subsyndromal attentional difficulties, mood and anxiety symptoms. Further, WASH-U-KSADS does not screen for ASD and there were no screening measures for ASD incorporated into project protocol. The project team did exclude 3 cases with ASD but these were based on clinical diagnoses made by the children's treating team.

Tables

Table 1: Group characteristics

	OBP	OHC
Parents		
	(n=17)	(n=23)
M:F ratio	1:16	-
Ethnicity	White British: 16 Asian Indian: 1	White British: 23
Socioeconomic status (Hollingshead and Redlich Scale)		
1	1 (5.88%)	1 (4.34%)
2	5 (29.41%)	8 (34.78%)
3	9 (52.94%)	11 (47.83%)
4	2 (11.77%)	3 (13.05%)
Offspring Characteristics		
	(n=24)	(n=34)
M:F	14:10	21:13
Age Range (months)	76-178	76-179
Mean age+SD (months)	141.13+31.89	124.18+28.92
Mean Full Scale IQ+SD	92.8+11.7	106.9+16.6

Table 2 Mixed-effects regression model examining the effect of predictor variables on the combined numbers of errors in labeling facial emotions

DANVA 2 Total Errors: 4 emotions combined				
Parameter	Value	Std. Error	t value	p value
Intercept	0.721	0.104	6.935	<0.00001
Parental BD	0.208	0.106	2.09	0.05*
SES	0.017	0.076	0.221	0.827
IQ	-0.001	0.003	-0.313	0.759
Age	-0.005	0.002	-3.452	0.004*
Male gender	0.103	0.1	1.029	0.321
Child faces	-0.167	0.047	-3.542	0.0004*
Low Intensity	0.405	0.042	9.623	<0.00001*

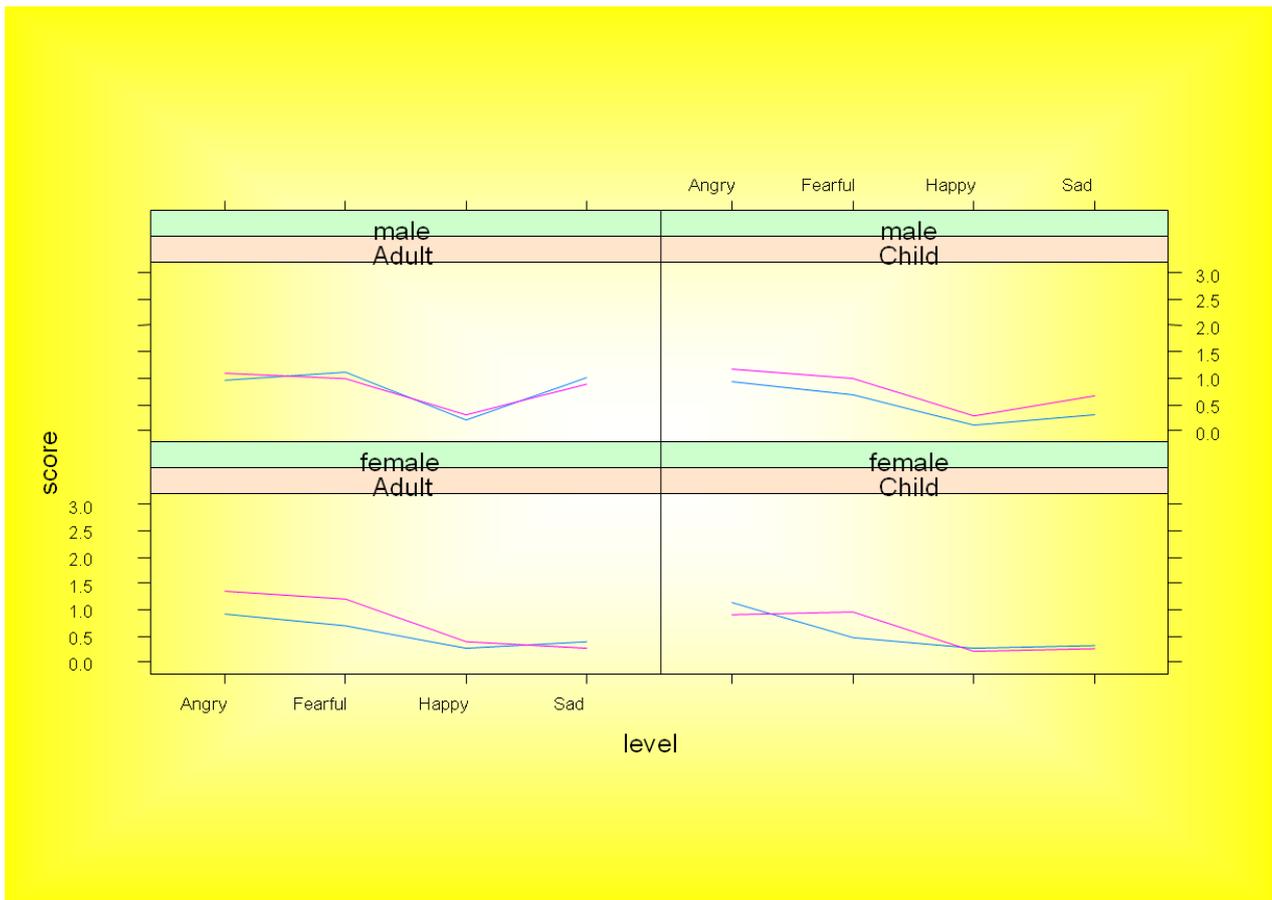
* indicates statistical significance

Table 3 Mixed-effects regression model examining the effect of predictor variables on the numbers of errors in labeling specific facial emotions

DANVA 2 Errors by Emotion				
Parameter	Value	Std. Error	t value	p value
Angry				
Intercept	0.352	0.143	2.465	0.015
Parental BD	0.239	0.159	1.501	0.142
SES	-0.004	0.002	-1.716	0.108
IQ	0.001	0.004	0.168	0.869
Age	0.109	0.114	0.954	0.346
Male gender	0.049	0.15	0.324	0.751
Child faces	0.064	0.101	0.631	0.529
Low intensity	0.966	0.09	10.7	<0.00001*
Sad				
Intercept	0.520	0.147	3.541	0.0005
Parental BD	0.1	0.156	0.643	0.524
SES	-0.074	0.111	-0.662	0.512
IQ	-0.003	0.004	-0.878	0.395
Age	-0.006	0.002	-2.435	0.029*
Male gender	0.313	0.146	2.140	0.05*
Child faces	-0.326	0.09	-3.612	0.0004*
Low intensity	-0.071	0.086	-0.823	0.412
Fearful				
Intercept	0.497	0.167	2.974	0.003
Parental BD	0.427	0.186	2.293	0.028*
SES	-0.008	0.003	-2.727	0.016*
IQ	0.003	0.005	0.613	0.55
Age	0.096	0.133	0.724	0.474*
Male gender	0.202	0.175	1.152	0.269
Child faces	-0.345	0.09	-3.825	<0.00001*
Low intensity	0.46	0.1	4.56	<0.00001*
Happy				
Intercept	0.029	0.064	0.454	0.651
Parental BD	0.116	0.074	1.569	0.125
SES	-0.003	0.001	-2.983	0.01
IQ	0.000	0.002	-0.187	0.855
Age	-0.024	0.053	-0.448	0.657
Male gender	-0.045	0.069	-0.645	0.53
Child faces	0.03	0.046	0.647	0.518
Low intensity	0.228	0.045	5.045	<0.00001*

* indicates statistical significance

Figure 1: errors made by OBP (pink) vs OHC (blue) by gender x angry, fearful, happy and sad facial emotion stimuli



Y axis: Number of errors

Bibliography

- American Psychiatric Association, 2013. Diagnostic and Statistical Manual of Mental Disorders (Fifth ed.). American Psychiatric Publishing, Arlington VA.
- Baum, K.M., M.C., L., Walker, E.F., Tomlinson, H., Schiffman, J., 1996. Emotion Recognition in adolescents with schizotypal personality disorder., Annual meeting of the Society for Research in Psychopathology, Atlanta, GA.
- Blair, R.J., Colledge, E., Murray, L., Mitchell, D.G., 2001. A selective impairment in the processing of sad and fearful expressions in children with psychopathic tendencies. *J Abnorm Child Psychol* 29, 491-498.
- Boyatzis, C.J., Chazan, E., Ting, C.Z., 1993. Preschool children's decoding of facial emotions. *J Genet Psychol* 154, 375-382.
- Bozikas, V.P., Tonia, T., Fokas, K., Karavatos, A., Kosmidis, M.H., 2006. Impaired emotion processing in remitted patients with bipolar disorder. *J Affect Disord* 91, 53-56.
- Brotman, M.A., Guyer, A.E., Lawson, E.S., Horsey, S.E., Rich, B.A., Dickstein, D.P., Pine, D.S., Leibenluft, E., 2008a. Facial Emotion Labeling Deficits in Children and Adolescents at Risk for Bipolar Disorder. *American Journal of Psychiatry* 165 (3), 385-389.
- Brotman, M.A., Skup, M., Rich, B.A., Blair, K.S., Pine, D.S., Blair, J.R., Leibenluft, E., 2008b. Risk for bipolar disorder is associated with face-processing deficits across emotions. *J Am Acad Child Adolesc Psychiatry* 47, 1455-1461.
- Craddock, N., Jones, I., 1999. Genetics of bipolar disorder. *J Med Genet* 36, 585-594.
- De Sonneville, L.M.J., Verschoor, C.A., Njokiktjien, C., Op het Veld, V., Toorenaar, N., Vranken, M., 2002. Facial identity and facial emotions: speed, accuracy, and processing strategies in children and adults. *J Clin Exp Neuropsychol* 24, 200-213.
- Dickstein, D.P., Brazel, A.C., Goldberg, L.D., Hunt, J.I., 2009. Affect regulation in pediatric bipolar disorder. *Child Adolesc Psychiatr Clin N Am* 18, 405-420, ix.
- Egger, H.L., Pine, D.S., Nelson, E., Leibenluft, E., Ernst, M., Towbin, K.E., Angold, A., 2011. The NIMH Child Emotional Faces Picture Set (NIMH-ChEFS): a new set of children's facial emotion stimuli. *Int J Methods Psychiatr Res* 20, 145-156.
- Ekman, P., Friesen, W.V., 1976. Pictures of facial affect. Consulting Psychologists Press, Palo Alto CA.
- First, M.B., Spitzer, R.L., Gibbon, M., Williams, J.B.W., 2002. Structured Clinical Interview for DSM-IV TR Axis I Disorders-Patient Edition (SCID-I/P). New York Psychiatric Institute, Biometric Research, New York.
- Fulford, D., Peckham, A.D., Johnson, K., Johnson, S.L., 2014. Emotion perception and quality of life in bipolar I disorder. *J Affect Disord* 152-154, 491-497.
- Geller, B., Williams, M., Zimmerman, B., Frazier, J., 1996. Washington University in St. Louis Kiddie Schedule for Affective Disorders and Schizophrenia (WASH-U-KSADS), St Louis, Washington University.
- George, M.S., Huggins, T., McDermet, W., Parekh, P.I., Rubinow, D., Post, R.M., 1998. Abnormal facial emotion recognition in depression: serial testing in an ultra-rapid-cycling patient. *Behav Modif* 22, 192-204.
- Gershon, E.S., Hamovit, J., Guroff, J.J., Dibble, E., Leckman, J.F., Sceery, W., Targum, S.D., Nurnberger, J.I., Jr., Goldin, L.R., Bunney, W.E., Jr., 1982. A family study of schizoaffective, bipolar I, bipolar II, unipolar, and normal control probands. *Arch Gen Psychiatry* 39, 1157-1167.
- Gershon, E.S., Mark, A., Cohen, N., Belizon, N., Braron, M., Knobe, K.E., 1975. Transmitted factors in the morbid risk of affective disorders: a controlled study. *J Psychiatr Res* 12, 283-299.
- Hanford, L.C., Sassi, R.B., Hall, G.B., 2016. Accuracy of emotion labeling in children of parents diagnosed with bipolar disorder. *J Affect Disord* 194, 226-233.
- Hollingshead, A.B., Redlich, F.C., 2007. Social class and mental illness: a community study. 1958. *Am J Public Health* 97, 1756-1757.
- Hsieh, F.Y., Lavori, P.W., Cohen, H.J., Feussner, J.R., 2003. An overview of variance inflation factors for sample-size calculation. *Eval Health Prof* 26, 239-257.
- Judd, L.L., Akiskal, H.S., Schettler, P.J., Endicott, J., Leon, A.C., Solomon, D.A., Coryell, W., Maser, J.D., Keller, M.B., 2005. Psychosocial disability in the course of bipolar I and II disorders: a prospective, comparative, longitudinal study. *Arch Gen Psychiatry* 62, 1322-1330.

Kaufman, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P., Williamson, D., Ryan, N., 1997. Schedule for Affective Disorders and Schizophrenia for School Age Children- Present and Lifetime Version (K-SADS-PL). *Journal of American Academy of Child and Adolescent Psychiatry* 36, 980-988.

Korkman, M., Kirk, U., Kemp, S.L., 2007. NEPSY II. Clinical and Interpretative Manual. Psychological Corporation, San Antonio, TC.

Lahera, G., Herrera, S., Reinares, M., Benito, A., Rullas, M., Gonzalez-Cases, J., Vieta, E., 2015. Hostile attributions in bipolar disorder and schizophrenia contribute to poor social functioning. *Acta Psychiatr Scand* 131, 472-482.

Leibenluft, E., Charney, D.S., Towbin, K.E., Bhangoo, R.K., Pine, D.S., 2003. Defining clinical phenotypes of juvenile mania. *Am J Psychiatry* 160, 430-437.

Lembke, A., Ketter, T.A., 2002. Impaired recognition of facial emotion in mania. *Am J Psychiatry* 159, 302-304.

Markham, R., Adams, K., 1992. The effect of type of task on childrens' identification of facial expressions. *Journal of Nonverbal Behavior* 16, 21-39.

McClure, E.B., 2000. A meta-analytic review of sex differences in facial expression processing and their development in infants, children, and adolescents. *Psychol Bull* 126, 424-453.

McClure, E.B., Pope, K., Hoberman, A.J., Pine, D.S., Leibenluft, E., 2003. Facial Expression Recognition in Adolescents with Mood and Anxiety Disorders. *American Journal of Psychiatry* 160, 1172-1174.

Mufson, L., Nowicki, S., Jr., 1991. Factors affecting the accuracy of facial affect recognition. *J Soc Psychol* 131, 815-822.

Nowicki, S., Carton, J., 1993. The measurement of emotional intensity from facial expressions: The DANVA faces 2. *Journal of Social Psychology* 133, 749-750.

Nowicki, S., Jr., Mitchell, J., 1998. Accuracy in identifying affect in child and adult faces and voices and social competence in preschool children. *Genet Soc Gen Psychol Monogr* 124, 39-59.

Olsavsky, A.K., Brotman, M.A., Rutenberg, J.G., Muhrer, E.J., Deveney, C.M., Fromm, S.J., Towbin, K., Pine, D.S., Leibenluft, E., 2012. Amygdala hyperactivation during face emotion processing in unaffected youth at risk for bipolar disorder. *J Am Acad Child Adolesc Psychiatry* 51, 294-303.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., R Core Team, 2016. nlme: Linear and NonLinear Mixed Effects Model. R package version 3.1 - 128, <http://CRAN.R-project.org/package=nlme>.

R Core Team, 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rich, B.A., Grimley, M.E., Schmajuk, M., Blair, K.S., Blair, R.J., Leibenluft, E., 2008. Face emotion labeling deficits in children with bipolar disorder and severe mood dysregulation. *Dev Psychopathol* 20, 529-546.

Rich, B.A., Vinton, D.T., Roberson-Nay, R., Hommer, R.E., Berghorst, L.H., McClure, E.B., Fromm, S.J., Pine, D.S., Leibenluft, E., 2006. Limbic hyperactivation during processing of neutral facial expressions in children with bipolar disorder. *Proc Natl Acad Sci U S A* 103, 8900-8905.

Roberts, G., Green, M.J., Breakspear, M., McCormack, C., Frankland, A., Wright, A., Levy, F., Lenroot, R., Chan, H.N., Mitchell, P.B., 2012. Reduced Inferior Frontal Gyrus Activation During Response Inhibition to Emotional Stimuli in Youth at High Risk of Bipolar Disorder. *Biol Psychiatry*.

Schepman, K., Taylor, E., Collishaw, S., Fombonne, E., 2012. Face emotion processing in depressed children and adolescents with and without comorbid conduct disorder. *J Abnorm Child Psychol* 40, 583-593.

Simonsen, C., Sundet, K., Vaskinn, A., Birkenaes, A.B., Engh, J.A., Hansen, C.F., Jonsdottir, H., Ringen, P.A., Opjordsmoen, S., Friis, S., Andreassen, O.A., 2008. Neurocognitive profiles in bipolar I and bipolar II disorder: differences in pattern and magnitude of dysfunction. *Bipolar Disord* 10, 245-255.

Torrent, C., Martinez-Aran, A., Daban, C., Sanchez-Moreno, J., Comes, M., Goikolea, J.M., Salamero, M., Vieta, E., 2006. Cognitive impairment in bipolar II disorder. *Br J Psychiatry* 189, 254-259.

Venn, H.R., Gray, J.M., Montagne, B., Murray, L.K., Michael Burt, D., Frigerio, E., Perrett, D.I., Young, A.H., 2004. Perception of facial expressions of emotion in bipolar disorder. *Bipolar Disord* 6, 286-293.

Wechsler, D., 1999. Wechsler Abbreviated Scale of Intelligence. The Psychological Corporation.

Wegbreit, E., Cushman, G.K., Puzia, M.E., Weissman, A.B., Kim, K.L., Laird, A.R., Dickstein, D.P., 2014. Developmental meta-analyses of the functional neural correlates of bipolar disorder. *JAMA Psychiatry* 71, 926-935.

Wegbreit, E., Weissman, A.B., Cushman, G.K., Puzia, M.E., Kim, K.L., Leibenluft, E., Dickstein, D.P., 2015. Facial emotion recognition in childhood-onset bipolar I disorder: an evaluation of developmental differences between youths and adults. *Bipolar Disord* 17, 471-485.

Whitney, J., Howe, M., Shoemaker, V., Li, S., Marie Sanders, E., Dijamco, C., Acquaye, T., Phillips, J., Singh, M., Chang, K., 2013. Socio-emotional processing and functioning of youth at high risk for bipolar disorder. *J Affect Disord* 148, 112-117.

Wingo, A.P., Baldessarini, R.J., Compton, M.T., Harvey, P.D., 2010. Correlates of recovery of social functioning in types I and II bipolar disorder patients. *Psychiatry Res* 177, 131-134.

Yurgelun-Todd, D.A., Gruber, S.A., Kanayama, G., Killgore, W.D.S., Baird, A.A., Young, A.D., 2000. fMRI during affect discrimination in bipolar affective disorder. *Bipolar Disord* 2, 237-248.