

# Let's Face It: Facial Emotion Processing Is Impaired in Bipolar Disorder

Tamsyn Elizabeth Van Rheenen,<sup>1,2</sup> AND Susan Lee Rossell<sup>1,2</sup>

<sup>1</sup>Brain and Psychological Sciences Research Centre, Swinburne University, Melbourne, Victoria, Australia

<sup>2</sup>Monash Alfred Psychiatry Research Centre, Central Clinical School, Monash University and the Alfred Hospital, Melbourne, Victoria, Australia

(RECEIVED July 22, 2013; FINAL REVISION November 24, 2013; ACCEPTED November 25, 2013; FIRST PUBLISHED ONLINE January 15, 2014)

## Abstract

Patients with bipolar disorder (BD) have difficulty in recognizing and discriminating facial emotions. However, beyond this broad finding, existing literature is equivocal about the specific nature of impairments, and progress toward adequately profiling facial emotion processing in BD is hampered by methodological inconsistencies. The current study aimed to advance the literature by comparing 50 BD patients and 52 controls on a series of facial emotion processing tasks. Results indicated that patients with BD had a small, yet consistent impairment in emotion processing overall. This impairment did not vary as a function of specific emotions, tasks, or intensities between groups, and was not influenced by current mood state. These results suggest that past inconsistencies in the literature are unlikely to be attributable to task related artifacts influencing the estimation of an effect. These findings add to our understanding of social cognition in BD, and have important implications for clinicians treating patients with the disorder. (*JINS*, 2014, *20*, 200–208)

**Keywords:** Cognition, Emotion, Neuropsychology, Emotion perception, Social cognition, Mood disorders, Emotion recognition

## INTRODUCTION

Perception of emotion from facial information is vital for effective social and relational functioning; misinterpretation of emotional expressions can lead to uncomfortable social situations and reduce appropriate social communication. There is growing evidence that impaired facial emotion processing ability is a feature of the social cognitive profile of bipolar disorder (BD; see Van Rheenen & Rossell, 2013b). This impairment is likely to be a factor in the problematic psychosocial functioning and reduced quality of life seen in BD (Hoertnagl et al., 2011; Martino, Strejilevich, Fassi, Marengo, & Igoa, 2011).

Studies investigating emotion processing have shown that patients with BD have an impaired capacity to recognize and discriminate facial emotions. This effect has been demonstrated in both symptomatic and euthymic samples as well as in at risk groups (e.g., Bozikas, Tonia, Fokas, Karavatos, & Kosmidis, 2006; Brotman, Guyer, et al., 2008; Getz, Shear, & Strakowski, 2003; Lembke & Ketter, 2002;

Vederman et al., 2012). Although this body of research is growing, there is still little clarity around the nature of the deficit. For example, although many studies have reported a general reduction in emotion processing accuracy (Brotman, Skup, et al., 2008; Derntl, Seidel, Kryspin-Exner, Hasmann, & Dobmeier, 2009; Getz, et al., 2003; Gray et al., 2006; Guyer et al., 2007), there are findings suggesting that abnormalities are more heavily weighted toward the processing of fear (Lembke & Ketter, 2002; Vederman et al., 2012), sadness (Derntl et al., 2009; Schenkel, Pavuluri, Herbener, Harral, & Sweeney, 2007; Vederman, et al., 2012), or surprise (Summers, Papadopoulou, Bruno, Cipolotti, & Ron, 2006). Some studies have found no deficit (Vaskinn et al., 2007), and some have found impairments of emotion discrimination but not labeling (Addington & Addington, 1998; Rossell, Van Rheenen, Groot, Gogos, & Joshua, 2013).

Progress toward adequately profiling facial emotion processing abilities in BD is currently hampered by several factors. First, existing studies differ in terms of task stimuli, with some using still photographic stimuli (static tasks), and some using morphing facial expressions (dynamic tasks). As there has been no study directly comparing performance across static and dynamic stimuli in the same BD cohort, the impact of subtle differences between task stimuli designs is

Correspondence and reprint requests to: Tamsyn Van Rheenen, Cognitive Neuropsychiatry Lab, Monash Alfred Psychiatry research centre (MAPrc), Level 4, 607 St Kilda Rd, Melbourne, VIC 3004, Australia. E-mail: tvanrheenen@swin.edu.au

not known. Dynamic tasks are arguably more ecologically valid; in the healthy population accuracy rates for facial emotion recognition are better for dynamic task designs which suggests that motion has a facilitatory effect on facial emotion perception (Ambadar, Schooler, & Cohn, 2005). Whether this effect extends to BD, however, remains to be seen.

Second, very few studies in BD have investigated both emotion labeling and discrimination performance. However, as these separate, albeit related abilities require different skills (Feinberg, Rifkin, Schaffer, & Walker, 1986; Walker, McGuire, & Bettes, 1984), it is uncertain as to whether facial emotion processing problems in BD reflect a specific difficulty in recognizing emotions on the basis of an impairment in matching emotional facial cues (discrimination), a specific difficulty in applying linguistic labels to different expressions (labeling), or a more generalized impairment involving both processes. Third, given the generally small effect sizes for facial emotion processing differences between patients and controls (Samamé, Martino, & Strejilevich, 2012; Vaskinn et al., 2007), it is likely that emotion processing impairments in BD are subtle. Therefore, slight alterations in the *intensity* of a facial expression stimulus could well influence observed group related differences in identification or discrimination of expressions. Although some studies have attempted to profile sensitivity thresholds for emotion in BD, there are few that have assessed the threshold of intensity at which emotions are most consistently identified. The former use paradigms in which respondents themselves alter the intensity of an expression until it reaches a level at which it is recognizable (Gray et al., 2006; Schaefer, Baumann, Rich, Luckenbaugh, & Zarate, 2010; Summers et al., 2006; Venn et al., 2004). However, these paradigms merely permit the assessment of how much intensity is required to perceive an emotion, rather than how reliably emotions are perceived at different intensities. To determine if BD is associated with reduced perceptual processing of emotional cues, research designs manipulating stimulus intensity are required. Demonstration that people with BD require higher levels of stimulus intensity to reliably label and discriminate facial emotions would constitute evidence for this hypothesized subtle but functionally important deficit.

In light of these factors, we set out to comprehensively examine facial emotion processing in a group of patients with BD compared to controls in a single experiment in which multiple variables were manipulated sequentially. Our objectives were fourfold in nature; first, we aimed to establish the comparability of two emotion labeling task stimuli designs (dynamic and static) by directly contrasting performance between them. Second, we aimed to test for group related differences in the consistency with which facial emotions are processed at different levels of intensity. Third, we aimed to determine the specificity of potential emotion processing deficits as a function of task type (labeling vs. discrimination). Finally, we aimed to determine whether emotion labeling performance varied as a function of emotion types between groups, to establish whether impairments generalize to a range of basic emotions, or are more weighted

toward a single emotion or subset of emotions in BD. Broadly we expected that BD-related impairments in facial emotion processing would be seen in both labeling and discrimination tasks. The following specific hypotheses were made; BD-related impairments will be seen irrespective of stimulus type (dynamic *versus* static: Hypothesis 1), BD-related impairments will be significant across all emotions (Hypothesis 2), BD-related impairments will be seen at all levels of stimulus intensity for both emotion labeling (Hypothesis 3a) and emotion discrimination (Hypothesis 3b), BD will not be associated with a deficiency in non-emotional identification of faces (Control task: Hypothesis 4).

## METHOD

This study was approved by the Alfred Hospital and Swinburne University Human Ethics Review Boards and abided by the Declaration of Helsinki. Written informed consent was obtained from each participant before the study began.

## Participants

The clinical sample comprised 50 patients (16 male, 34 female) diagnosed as having DSM-IV-TR BD (39 BD I, 12 BD II) using the Mini International Neuropsychiatric Interview (MINI: Sheehan et al., 1998). Patients were recruited *via* community support groups and general advertisements and were all out-patients. Current symptomology was assessed using the Young Mania Rating Scale (YMRS: Young, Biggs, Ziegler, & Meyer, 1978) and the Montgomery Asberg Depression Rating Scale (MADRS: Montgomery & Asberg, 1979); there were 17 depressed (defined as those that met strict criteria for MADRS scores > 8), 12 mixed (defined as those that met strict criteria for YMRS and MADRS scores > 8), 4 (hypo)manic (defined as those that met strict criteria for YMRS scores > 8) and 17 euthymic (defined as those that met strict criteria for YMRS and MADRS scores ≤ 8) patients (i.e., 33 that were symptomatic). Patients with current psychosis, co-morbid psychotic disorders, visual impairments, neurological disorder and/or a history of substance/alcohol abuse or dependence during the past six months were excluded. Thirty-two patients were taking antipsychotics, 16 were taking antidepressants, 16 were taking mood stabilizers and 10 were taking benzodiazepines.<sup>1</sup> Demographic and clinical characteristics are presented in Table 1.

A control sample of 52 healthy participants (20 male, 32 female) were recruited for comparison purposes by general advertisement and contacts of the authors. Using the MINI screen, no control participant had a current diagnosis or previous history of psychiatric illness (Axis I). An immediate family history of mood and psychiatric disorder, in addition to a personal history of neurological disorder, current or previous alcohol/substance dependence or abuse, visual impairments

<sup>1</sup> Repeated measures ANOVAs using medication (dichotomously coded to yes/no) as the between subjects factor revealed no significant difference on any of the tasks for patients on or off any of the classes of medication

and current psychiatric medication use was exclusion criteria for all controls.

All participants were fluent in English, were between the ages of 18 and 65 years and had an estimated pre-morbid IQ as scored by the Wechsler Test Of Adult Reading (WTAR) of >90.

## Materials

All participants completed three computerized tasks (designed by the author) that were used to measure (1) emotion labeling performance across stimulus type (dynamic *versus* static) and emotion type (happy, sad, angry, fear, neutral), (2) emotion discrimination, and (3) emotion labeling and discrimination performance across three levels of intensity. Participants also completed a control task (designed by the authors) to test whether potential impairments on the facial emotion processing tasks were reflective of a generalized performance deficit or were specific to facial emotion processing.

The face stimuli were taken from the widely used and well validated Ekman and Friesen series known as the Pictures of Facial Affect (POFA: Ekman & Friesen, 1976). The stimuli comprised black and white photographs of faces free of jewelry, spectacles, make up and facial hair (five female and five male) and expressing the emotions happy, sad, fear, angry and neutral. The faces were cropped to an oval shape spanning the top of the forehead to the bottom of the chin and excluding any hair and the ears on either side of the face. Thirteen faces were used in total. However, given that we endeavored to have an equal presentation of emotional expressions of different genders in the labeling tasks across ten different trials for each emotion and neutral, and that some expressions were not available for all faces, the tasks presented below varied in the exact faces used.

A morphing program called Fantamorph (Abrosoft, 2012) was used to reduce the intensity of the POFA stimuli's emotional expression by 25% decrements to create static intensity varied stimuli. This resulted in static stimulus expressions at 100%, 75%, 50%, and 25% intensity. Pilot testing revealed that emotions presented at 25% intensity elicited floor effects in controls, and they were therefore excluded from the task. The final stimulus set for the latter two tasks comprised static faces displaying 100%, 75%, and 50% (high, medium, and low) emotional intensity only. The dynamic stimuli were created by morphing the low, medium, and high intensity static faces through quick successive frames from a neutral expression (0%) to the final emotional expression (100%), such that they appeared as a moving image. All tasks were presented on a 14" Lenovo laptop computer and were run through Presentation (Neurobehavioral Systems Inc, 2012). These tasks are described below.

*The dynamic facial emotion labeling task* required participants to view dynamic facial images (i.e., morphs) and identify the emotion being expressed. Forty randomized dynamic display trials comprising 10 presentations (5 male and 5 female faces) each for happy, sad, fear, and angry expressions

were presented one at a time for 1500 ms<sup>2</sup> followed by an inter-stimulus interval of 1500 ms. Participants were instructed to press a labeled keyboard button corresponding to the emotion that they believed the face was expressing as soon as they recognized it. The averaged accuracy percentage for each emotion was taken as the primary dependent variable. This task was used in the analyses testing Hypothesis 1 and 2.

*The static facial emotion labeling task* was designed to assess participants' ability to identify emotional expressions from static facial stimuli (i.e., a photograph). It required participants to view an image of a male or female face, and identify the emotional expression exhibited by that face. The faces were presented one at a time on a black background for 2000 ms followed by an inter-stimulus interval of 1500 ms. The task involves 130 randomized trials in total, including 10 presentations (5 male and 5 female faces) for each of the emotions happy, sad, angry, and fear at each of the three levels of emotional intensity (high, medium, and low) and 10 presentations of neutral. Participants were instructed to press a labeled keyboard button corresponding to the emotion that they believed the face was expressing as soon as they recognized it. The averaged accuracy percentage and response time for each emotion at each level of intensity was taken as the primary dependent variable. This task was used in analyses testing Hypotheses 1 and 3a.

*The static facial emotion discrimination task* was designed to assess participants' ability to differentiate between static facial emotional expressions at different levels of intensity. It required participants to view two simultaneously presented images of human faces and identify whether the emotion that the two faces were showing was the same or different. The task represents the emotions happy, sad, angry, fear, and neutral over 135 randomized paired stimulus trials at the three different intensities.<sup>3</sup> Emotional expressions were paired only with those expressing the same level of intensity (i.e., high intensity expressions paired together) or a neutral expression, but never with an emotion of a different expressive intensity. There were 31 incongruent paired trials (representing pairings across the emotions happy, sad, angry, fear, and neutral) and 11 congruent paired trials (representing pairings across happy, sad, angry, and fear) for each level of intensity, with 9 additional trials representing paired neutral expressions that were used as fillers. One face in each pair was presented to the left visual field, and the other was presented to the right visual field on a black background for 2000 ms followed by an inter-stimulus interval of 1500 ms. Participants were instructed to respond *via* two button keyboard press (same or different) as soon as they could

<sup>2</sup> Emotional morphs were trialed at different presentation durations. It was decided that 1500 ms was the most realistic timeframe representation of a developing emotional expression and this time frame was used to maintain ecological validity.

<sup>3</sup> It should be noted that the number of stimuli in each set differed depending on the nature of the task. For the two labeling tasks, there were 10 presentations of each emotional expression. In the static task, this occurred per intensity in addition to 10 presentations of neutral. As neutral morphs and different intensities were not available in the dynamic task, the number of trials in that task was reduced (in comparison to the static task).

discriminate the emotional expressions. Responses made from 200 ms onward were recorded. Written instructions, an example and a set of practice trials were provided to participants before commencing the task. The averaged accuracy percentage and response times across congruent and incongruent trials at each level of intensity was taken as the primary dependent variable for this task. This task was used in analyses testing Hypothesis 3b.

To rule out the possibility that BD is associated with a more fundamental deficit in facial processing (i.e., not specific to higher order facial emotion processing), a *static identity discrimination task* was designed. The tasks assessed participant's ability to determine whether two simultaneously presented static facial stimuli were identical or not. The task comprised 55 (45 incongruent and 10 congruent) randomized trials made up of couplings between six male and six female faces displaying neutral expressions only. One face in each pair was presented to the left visual field, and the other was presented to the right visual field on a black background for 2000 ms followed by an inter-stimulus interval of 1500 ms. The task involved 65 randomized trials with the same pairs being presented twice. Participants were instructed to respond *via* two button keyboard press (same or different) as soon as they could discriminate the faces. Responses made from 200 ms onward were recorded with the averaged accuracy percentage and response times across same/different pairs taken as the dependent variables. This task was used in the analysis testing Hypothesis 4.

## Statistical Analysis

Demographic and clinical group differences were assessed *via* independent samples *t* tests or  $\chi^2$  tests. We conducted a series of analyses to address our hypotheses: we used a four (emotion: happy, sad, angry, fear) \* two (stimuli type: static, dynamic) \* two (group: control, BD) repeated measures analysis of variance (ANOVA) of the accuracy data to address Hypothesis 1 and 2. Significant main effects of group, emotion, and of stimuli type were expected to show support for this hypothesis. Only responses to high intensity conditions were used as the static emotion labeling variables in this analysis. Due to the response time windows differing between dynamic and static stimuli tasks, we were unable to analyze response time differences across these tasks.<sup>4</sup> Two repeated measures ANOVAs using a four (emotion: happy, sad, angry, fear) \* three (intensity: high, medium, low) \* two (group: control, BD) design with the accuracy and response

<sup>4</sup> Accuracy performance for the labeling of happy emotions was found to be at ceiling for both groups, so we re-ran the analysis excluding happy. Results remained largely unchanged, so only the analyses including happy are reported here. To determine whether having neutral anchoring expressions as a point of comparison impacted static facial emotion labeling effects, we re-ran the repeated measures ANOVA including neutral expressions and the four emotions used in the previous analysis in a five (emotions: happy, sad, angry, fear, and neutral) \* two (group: controls, BD) design with only the static emotion labeling data. This analysis obviously did not include the static/dynamic within-group contrast. However, as this analysis made no difference to the participant \* group interaction or between-group effects, again for brevity it is not presented.

time data of the static emotion labeling task were used to address Hypothesis 3a. To address Hypothesis 3b we also completed two, three (intensity: high, medium, low) \* two (group: control, BD) repeated measures ANOVAs on the static emotion discrimination task, separately for the accuracy and response time data. Support for these hypotheses was expected to be shown through main effects of group and intensity. Two one-way ANOVAs were used to compare accuracy and response time performance between groups on the control task (i.e., the static identity discrimination task), and thus to address Hypothesis 4. The absence of a significant group effects would demonstrate support for this hypothesis.

To better understand the effects of mood and diagnostic status on facial emotion processing performance, all analyses were re-run in the patient group: in the first series of analyses, diagnosis (BD I;  $n = 38$  or BD II;  $n = 12$ ) was entered as the between-groups factor for all tasks. For the second series of analyses, current mood state was entered as the between-groups factor; however, given that the sample size of some of the mood state subgroups was too small for meaningful analysis, we collapsed the mixed ( $n = 12$ ) and manic ( $n = 4$ ) groups into one (resulting  $n = 16$ ) and compared this to patients meeting criteria for euthymia ( $n = 17$ ) or depression ( $n = 17$ ). Bivariate correlations were also conducted to examine the relationship between emotion labeling and discrimination performance and symptom severity on the YMRS and MADRS. All analyses were corrected for multiple testing using a conservative  $\alpha$  set at .01.

## RESULTS

No significant differences in age, gender, education level completed, or pre-morbid IQ were found between the two groups (see Table 1).

### Emotion Labeling Accuracy as a Function of Task Stimuli Type (Hypotheses 1 and 2)

Investigation of the determinants of emotion labeling accuracy found main effects of emotion (Greenhouse Geisser corrected  $F(2,715, 271.538) = 71.51$ ;  $p < .001$ , *partial*  $\eta^2 = .42$ ), stimuli type: static *versus* dynamic ( $F(1,100) = 8.21$ ;  $p < .01$ ; *partial*  $\eta^2 = .08$ ) and group ( $F(1,100) = 7.04$ ;  $p < .01$ ; *partial*  $\eta^2 = .07$ ), but no two- or three-way interactions reached significance. Accuracy across emotions occurred in the descending order of happy ( $M = 98.43$ ;  $SD = 2.90$ ), fear ( $M = 86.32$ ;  $SD = 12.60$ ), angry ( $M = 81.76$ ;  $SD = 13.67$ ), and sad ( $M = 79.59$ ;  $SD = 13.93$ ), with performance being slightly better for the dynamic stimuli task relative to the static stimuli task (dynamic:  $M = 87.60$ ;  $SD = 8.02$ ; static  $M = 85.44$ ;  $SD = 9.20$ ;  $d = -0.25$ ) and BD patients performing less accurately than controls overall (Control:  $M = 88.43$ ;  $SD = 6.73$ ; BD:  $M = 84.50$ ;  $SD = 8.30$ ;  $d = -0.52$ ). Figure 1 presents emotion labeling accuracy performance for dynamic and static stimuli tasks as a function of emotion across groups.



**Table 1.** Demographic and clinical characteristics of the sample

Group	N	Control		N	BD		Group comparisons	
		M	SD		M	SD	F/ $\chi^2$	p
N	52			50				
Age		33.98	14.27		37.92	12.45	-1.48	.14
Gender (M/F)	20/32			16/34			0.47 <sup>^</sup>	.50
WTAR (scaled)		111.65	7.24		109.40	11.95	.96	.34
Education standard completed							8.70 <sup>^</sup>	.12
Completed secondary	12	-	-	7	-	-	-	-
Completed TAFE/diploma	3	-	-	11	-	-	-	-
Completed trade qualification	3	-	-	4	-	-	-	-
Completed tertiary degree	28	-	-	20	-	-	-	-
Other	4	-	-	8	-	-	-	-
Current mood state ([Hypo]manic/ Depressed/ Mixed/ Euthymic)		-	-	4/17/12/17	-	-	-	-
Diagnostic subtype (BD I/BD II)		-	-	38/12	-	-	-	-
Age of onset		-	-		20.49	8.45	-	-
Age of diagnosis		-	-		27.40	9.99	-	-
YMRS		-	-		6.33	5.48	-	-
MADRS		-	-		11.96	10.07	-	-

Note. <sup>^</sup> Group comparisons all independent samples t-tests except gender and education which was chi-squared; M/F = Male/Female, WTAR = Wechsler Test of Adult Intelligence, YMRS = Young Mania Rating Scale, MADRS = Montgomery Asberg Depression Rating Scale.

### Static Emotion Labeling as a Function of Emotion and Intensity (Hypothesis 3a)

Table 2 presents the accuracy and response time means and standard deviations for static emotion labeling as a function of intensity across groups. Analyses were conducted separately for the two dependent variables.

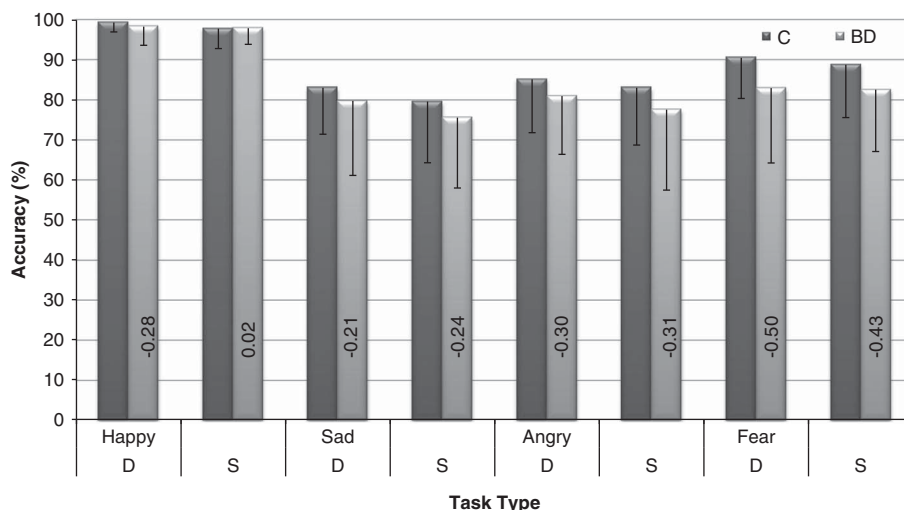
#### Accuracy

There was a main effect of emotion ( $F(3,297) = 152.37$ ;  $p < .001$ , *partial*  $\eta^2 = .61$ ) and intensity (Greenhouse Geisser corrected  $F(1.86, 184.22) = 621.83$ ;  $p < .001$ , *partial*  $\eta^2 = .86$ ), but no group effect or two- or three-way interactions. Accuracy across emotions occurred in the descending order of

happy ( $M = 91.74$ ;  $SD = 6.57$ ), fear ( $M = 75.52$ ;  $SD = 13.82$ ), sad ( $M = 69.51$ ;  $SD = 14.94$ ) and angry ( $M = 60.98$ ;  $SD = 14.05$ ), with performance best in high intensity ( $M = 85.44$ ;  $SD = 9.20$ ), followed by medium intensity ( $M = 80.02$ ;  $SD = 9.52$ ) and low intensity ( $M = 57.72$ ;  $SD = 11.62$ ) conditions. Although not significant, BD patients performed less accurately than controls and the effect size difference was in the medium range (Control:  $M = 76.01$ ;  $SD = 8.18$ ; BD:  $M = 72.73$ ;  $SD = 9.56$ ;  $d = -0.37$ ).

#### Response time

There was a main effect of emotion (Greenhouse Geisser corrected  $F(2.65, 265.30) = 190.03$ ;  $p < .001$ ;



**Fig. 1.** Emotion labeling accuracy performance for dynamic and static stimuli tasks as a function of emotion across groups.

**Table 2.** Accuracy and response time means and standard deviations for static emotion labelling as a function of intensity

Intensity	High					Medium					Low				
	Control		BD		<i>d</i>	Control		BD		<i>d</i>	Control		BD		<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Accuracy (%)															
Happy	97.88	4.98	98.00	4.040	0.02	96.47	6.88	95.40	8.38	-0.14	82.50	14.26	79.00	14.18	-0.25
Sad	79.62	15.34	75.60	17.63	-0.24	73.65	15.09	74.40	19.18	0.04	58.85	16.65	54.80	21.78	-0.21
Angry	83.08	14.35	77.60	20.16	-0.31	69.81	19.56	64.20	19.07	-0.29	37.12	18.61	33.80	15.50	-0.19
Fear	88.85	13.23	82.60	15.50	-0.43	86.54	12.35	79.40	16.59	-0.49	57.50	18.77	58.00	23.82	0.02
Overall	87.36	8.55	83.45	9.51	-0.43	81.65	7.91	78.35	10.75	-0.35	58.99	10.87	56.40	12.32	-0.22
Response time (ms)															
Happy	1069.73	207.25	1207.77	226.04	0.64	1111.25	207.65	1251.83	247.80	0.61	1322.59	249.86	1448.59	253.11	0.50
Sad	1364.69	257.57	1493.44	261.81	0.50	1404.48	238.00	1520.97	271.97	0.46	1500.48	281.90	1549.53	265.83	0.18
Angry	1424.85	345.78	1486.64	259.48	0.20	1483.81	323.56	1537.65	259.34	0.18	1530.10	465.50	1677.75	298.19	0.38
Fear	1480.36	268.08	1566.87	242.83	0.34	1553.42	274.06	1633.97	257.58	0.30	1718.86	282.62	1757.55	261.55	0.14
Overall	1334.91	231.18	1438.68	217.26	0.46	1388.24	218.44	1486.10	215.80	0.45	1518.01	224.56	1608.36	216.30	0.41

Note: BD = bipolar disorder; *d* = Cohen's *d*; static labelling group accuracy effect *p* = .07; static labelling group response time effect *p* < .05.

partial  $\eta^2 = .66$ ), intensity (Greenhouse Geisser corrected  $F(1.62, 161.94) = 98.901$ ;  $p < .001$ ; partial  $\eta^2 = .50$ ), and group ( $F(1,100) = 5.61$ ;  $p < .05$ ; partial  $\eta^2 = .05$ ), although the latter did not survive statistical correction. There were no two- or three-way interactions. Response latencies across emotions occurred in the ascending order of happy ( $M = 1233.55$ ,  $SD = 224.86$ ), sad ( $M = 1471.30$ ,  $SD = 238.63$ ), angry ( $M = 1522.61$ ,  $SD = 245.00$ ), and fear ( $M = 1617.83$ ,  $SD = 235.61$ ), with performance best in high intensity ( $M = 1385.78$ ,  $SD = 229.36$ ), followed by medium intensity ( $M = 1436.21$ ,  $SD = 221.60$ ) and low intensity ( $M = 1562.30$ ,  $SD = 224.10$ ) conditions. Although not significant, BD patients had longer latencies than controls and the effect size difference was in the medium range (Control:  $M = 1413.72$ ,  $SD = 209.27$ ; BD:  $M = 1511.05$ ,  $SD = 205.63$ ;  $d = 0.47$ ).

**Static Emotion Discrimination as a Function of Intensity (Hypothesis 3b)**

Table 3 presents the accuracy and response time means and standard deviations for static emotion discrimination as a function of intensity across groups. Analyses were conducted separately for the two dependent variables.

**Accuracy**

There was a main effect of intensity ( $F(2,200) = 97.78$ ;  $p < .001$ ; partial  $\eta^2 = .49$ ) and group ( $F(1,100) = 9.04$ ;

$p < .01$ ; partial  $\eta^2 = .08$ ), but no two-way interaction. Performance accuracy was best in high intensity ( $M = 78.85$ ;  $SD = 8.89$ ) followed by medium intensity ( $M = 69.40$ ;  $SD = 10.88$ ) and low intensity ( $M = 65.59$ ;  $SD = 9.55$ ) conditions, with overall performance being worse in patients than controls (BD:  $M = 68.94$ ;  $SD = 7.94$ ; Controls  $M = 73.52$ ;  $SD = 7.43$ ;  $d = -0.60$ ).

**Response time**

There was a main effect of intensity ( $F(2,200) = 69.33$ ;  $p < .001$ ; partial  $\eta^2 = .41$ ), but no effect of group and no two-way interaction. Response latencies were shortest in medium intensity ( $M = 1544.58$ ;  $SD = 198.71$ ) followed by high intensity ( $M = 1598.50$ ;  $SD = 281.22$ ) and low intensity ( $M = 1700.73$ ;  $SD = 286.84$ ) conditions.

**Control Task Performance (Hypothesis 4)**

There were no significant accuracy or response time differences between groups on the identity labeling control task (both *p*'s > .05).

**Subgroup Analyses**

There were no between-group main effects or interactions on any of the tasks for patients diagnosed as having BD I versus BD II (all *p*'s > .05), nor were there any between-group

**Table 3.** Accuracy and response time means and standard deviations for static emotion discrimination as a function of intensity

Intensity	High					Medium					Low				
	Control		BD		<i>d</i>	Control		BD		<i>d</i>	Control		BD		<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Accuracy (%)	80.52	8.27	77.11	9.26	-0.39	72.18	10.39	66.50	10.71	0.54	67.87	8.95	63.22	9.65	-0.50
Response time (ms)	1578.45	286.88	1619.26	276.56	0.14	1535.68	198.03	1553.85	201.42	0.09	1706.64	289.67	1694.59	286.68	0.04

Note: BD = bipolar disorder; RT = response time; *d* = Cohen's *d*; static discrimination group accuracy effect  $p < .01$ ; static labelling group response time effect  $p = .75$ .

main effects or interactions on any of the tasks for patients classified as euthymic, depressed or mixed/manic (all  $p$ 's > .05). Furthermore, bivariate correlation analyses found no significant associations between accuracy or response time performance (where applicable) on any measure and severity of current depression (MADRS score) or mania (YMRS score).

## DISCUSSION

The current study examined facial emotion processing in a cohort of BD patients compared to controls in a complex experiment whereby multiple variables were manipulated. In addition to examining whether emotion labeling performance varied as a function of specific emotions between groups, we aimed to establish the comparability of two commonly used emotion labeling task stimuli designs, determine the existence of differences in facial emotion processing at varying levels of intensity, and establish the specificity of emotion processing deficits as a function of task types; labeling *versus* discrimination. These aims were formulated to further understandings of the emotion processing profile of BD.

Our results indicated that emotion labeling performance was better for the dynamic task stimuli relative to the static task stimuli, with patients performing worse overall (supporting Hypothesis 1). This is consistent with previous research in the healthy population indicating that emotion processing performance is more accurately assessed using tasks that enable the facilitatory effect of motion to guide emotion recognition (Ambadar et al., 2005). Furthermore, given the significant group and emotion effects in the absence of interactions between them, there was no evidence that BD-related impairments were more heavily weighted toward a particular emotion. Rather BD patients appear to be globally compromised in processing a range of emotional expressions (supporting Hypothesis 2). We also found that the accuracy with which facial emotions were labeled and discriminated diminished in line with the degradation of stimulus intensity across groups. However, BD patients did exhibit overall deficits in discriminating emotions across these intensities compared to their control counterparts (supporting Hypothesis 3b). Contrary to expectations, this group effect was not evident for intensity labeling performance (i.e., there was no support for Hypothesis 3a), although the effect size difference between groups was still in the medium range. Finally, as no group difference was observed on the control task, it appears that the general processing of facial information was not compromised in this cohort (supporting Hypothesis 4).

Taken together, this pattern of findings suggests that BD patients have a relatively generalized impairment for the labeling of facially conveyed emotional expressions, which cannot be attributed to a pervasive impairment in the general processing of faces. These results accord with several studies in which emotion labeling (Derntl et al., 2009; Getz et al., 2003; Vederman et al., 2012) but not face processing itself

(e.g., Bozikas et al., 2006; Getz et al., 2003) has been shown to be impaired in BD. Thus it appears that BD related facial emotion processing difficulties reflect inability in both adequately matching facial cues of emotion to recognize expressions and adequately applying or understanding the linguistic labels used to identify them. Moreover, as diagnostic subtype and current mood state did not have any influence on the present findings, the generalized facial emotion processing impairment we have observed here is likely to be reflective of a trait-like feature of the disorder, which is consistent with past research (Bozikas et al., 2006; Vederman et al., 2012).

Importantly, it appears that whilst variability between tasks designs and intensities may subtly influence the *strength* at which emotion processing abilities are apparent (i.e., performance accuracy is better at higher compared to lower levels of intensity, and in tasks using dynamic instead of static stimuli), these procedural factors are unlikely to significantly impact the detection of a facial emotion processing effect in BD (as evidenced by the main effects of group, but not interactions across analyses). Thus, it is improbable that inconsistencies evident in the current BD literature are by-products of emotion processing impairments being masked by task related artifacts (as BD performance on all emotion processing tasks was impaired here). Rather, they may be a function of other factors such as differences amongst study cohorts with regard to clinical history, or the use and dosage of medications. Alternatively, the null effects of past research may represent a result of poor statistical power (see Vaskinn et al., 2007 who's BD sample comprised only 21 patients).

The present results should be interpreted within the confines of several limitations. First, as emotion processing under time pressure relies on general processing speed which is known to be compromised in BD (see Van Rheenen & Rossell, 2013a), it is possible that the effects we have observed here are confounded by generic BD-related cognitive impairments. Second, given that there was overlap in the stimuli used across the different tasks, it is possible that our results are partly attributable to cross-contamination effects whereby responses on earlier trials affected responses on later trials using the same face and facial expression. Third, as abnormalities of disgust and surprised expressions have been demonstrated in some studies (Gray et al., 2006; Harmer, Grayson, & Goodwin, 2002), our omission of these emotions from the battery limited our understanding of how accurately and quickly patients in this cohort were able to process these emotions. Fourth, as we did not explicitly counterbalance the presentation of faces and emotions across visual fields, we cannot account for hemisphere specific laterality effects. Finally, we were unable to directly compare mood subgroups to controls due to the restricted power after stratification into mixed/manic, depressed and euthymic subgroups. Although within group analyses failed to differentiate performance across patients in these current states for all tasks, it is still possible that mood may have had an effect on performance. Thus, given the rather heterogeneous

nature of our BD sample, our results should be interpreted with caution.

Nevertheless, as this study is the first of its kind to provide insight into the nature of emotion processing impairments in BD while paying attention to a range of potential confounds including task stimuli designs, intensity, and emotion specific factors in a single experiment, these findings do add substantially to the existing literature on facial emotion processing in BD. Future studies would certainly do well to address the present limitations however, with a view to providing greater clarity with regard to the impact of pre-existing cognitive impairments, cross-contamination effects and mood state related factors on facial emotion processing.

In summary, this study is the first of its kind to comprehensively examine emotion processing performance in a battery that controlled for subtle differences in task stimuli, and investigated the specificity of impairments across task types, emotions, and intensities in persons with BD. Our primary finding of a generalized patient impairment in the ability to label facial expressions and to make use of available emotional facial cues to differentiate them, suggests that facial emotion processing is considerably more challenging for people with the disorder than for those without. This may have direct impact on the significant psychosocial burden carried by patients, although this remains to be seen.

## ACKNOWLEDGEMENTS AND CONFLICTS OF INTEREST

The authors have no conflicts of interest but would like to acknowledge the Australian Rotary Health/Bipolar Expedition, the Helen McPherson Smith Trust and an Australian Postgraduate Award for providing financial assistance for the completion of this work.

## REFERENCES

- Abrosoft (2012). Fantamorph.
- Addington, J., & Addington, D. (1998). Facial affect recognition and information processing in schizophrenia and bipolar disorder. *Schizophrenia Research*, *32*, 171–181.
- Ambadar, Z., Schooler, J.W., & Cohn, J.F. (2005). Deciphering the enigmatic face. *Psychological Science*, *16*, 403–410. doi:10.1111/j.0956-7976.2005.01548.x
- Bozikas, V.P., Tonia, T., Fokas, K., Karavatos, A., & Kosmidis, M.H. (2006). Impaired emotion processing in remitted patients with bipolar disorder. *Journal of Affective Disorders*, *91*, 53–56.
- Brotman, M.A., Guyer, A.E., Lawson, E.S., Horsey, S.E., Rich, B.A., Dickstein, D.P., ... Leibenluft, E. (2008). Facial emotion labeling deficits in children and adolescents at risk for bipolar disorder. *American Journal of Psychiatry*, *165*, 385–389.
- Brotman, M.A., Skup, M., Rich, B.A., Blair, K.S., Pine, D.S., Blair, J.S., & Leibenluft, E. (2008). Risk for bipolar disorder is associated with face processing deficits across emotions. *Journal of the American Academy of Child and Adolescent Psychiatry*, *47*, 1455–1461.
- Derntl, B., Seidel, E.-M., Kryspin-Exner, I., Hasmann, A., & Dobmeier, M. (2009). Facial emotion recognition in patients with bipolar I and bipolar II disorder. *British Journal of Clinical Psychology*, *48*, 363–375. doi:10.1348/014466509X404845
- Ekman, P., & Friesen, W.V. (1976). *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologists Press.
- Feinberg, T.E., Rifkin, A., Schaffer, C., & Walker, E. (1986). Facial discrimination and emotional recognition in schizophrenia and affective disorders. *Archives of General Psychiatry*, *43*, 276–279. doi:10.1001/archpsyc.1986.01800030094010
- Getz, G.E., Shear, P.K., & Strakowski, S.M. (2003). Facial affect recognition deficits in bipolar disorder. *Journal of the International Neuropsychological Society*, *9*, 623–632.
- Gray, J., Venn, H., Montagne, B., Murray, L., Burt, M., Frigerio, E., ... Young, A.H. (2006). Bipolar patients show mood-congruent biases in sensitivity to facial expressions of emotion when exhibiting depressed symptoms, but not when exhibiting manic symptoms. *Cognitive Neuropsychiatry*, *11*, 505–520. doi:10.1080/13546800544000028
- Guyer, A.E., McClure, E.B., Adler, A.D., Brotman, M.A., Rich, B.A., Kimes, A.S., ... Leibenluft, E. (2007). Specificity of facial expression labeling deficits in childhood psychopathology. *Journal of Child Psychology and Psychiatry*, *48*, 863–871. doi:10.1111/j.1469-7610.2007.01758.x
- Harmer, C.J., Grayson, L., & Goodwin, G.M. (2002). Enhanced recognition of disgust in bipolar illness. *Biological Psychiatry*, *51*, 298–304.
- Hoertnagl, C.M., Muehlbacher, M., Biedermann, F., Yalcin, N., Baumgartner, S., Schwitzer, G., ... Hofer, A. (2011). Facial emotion recognition and its relationship to subjective and functional outcomes in remitted patients with bipolar I disorder. *Bipolar Disorders*, *13*, 537–544. doi:10.1111/j.1399-5618.2011.00947.x
- Lembke, A., & Ketter, T.A. (2002). Impaired recognition of facial emotion in mania. *American Journal of Psychiatry*, *159*, 302–304.
- Martino, D.J., Strejilevich, S.A., Fassi, G., Marengo, E., & Igoa, A. (2011). Theory of mind and facial emotion recognition in euthymic bipolar I and bipolar II disorders. *Psychiatry Research*, *189*, 379–384. doi:10.1016/j.psychres.2011.04.033
- Montgomery, S.A., & Asberg, M. (1979). A new depression scale designed to be sensitive to change. *British Journal of Psychiatry*, *134*, 382–389.
- Neurobehavioral Systems Inc (2012). Presentation, Version 14.8.
- Rossell, S.L., Van Rheenen, T.E., Groot, C., Gogos, A., & Joshua, N.R. (2013). Investigating affective prosody in psychosis: A study using the Comprehensive Affective Testing System. *Psychiatry Research*, *210*, 896–900. doi:10.1016/j.psychres.2013.07.037
- Samamé, C., Martino, D.J., & Strejilevich, S.A. (2012). Social cognition in euthymic bipolar disorder: Systematic review and meta-analytic approach. *Acta Psychiatrica Scandinavica*, *125*, 266–280. doi:10.1111/j.1600-0447.2011.01808.x
- Schaefer, K.L., Baumann, J., Rich, B.A., Luckenbaugh, D.A., & Zarate, C.A. Jr (2010). Perception of facial emotion in adults with bipolar or unipolar depression and controls. *Journal of Psychiatric Research*, *44*, 1229–1235.
- Schenkel, L.S., Pavuluri, M.N., Herbener, E.S., Harral, E.M., & Sweeney, J.A. (2007). Facial emotion processing in acutely ill and euthymic patients with pediatric bipolar disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, *46*, 1070–1079.
- Sheehan, D.V., Lecrubier, Y., Harnett Sheehan, K., Amorim, P., Janavs, J., Weiller, E., ... Dunbar, G.C. (1998). The Mini-International Neuropsychiatric Interview (MINI): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *The Journal of Clinical Psychiatry*, *59*, 22–33.



- Summers, M., Papadopoulou, K., Bruno, S., Cipolotti, L., & Ron, M.A. (2006). Bipolar I and bipolar II disorder: Cognition and emotion processing. *Psychological Medicine, 36*, 1799–1809.
- Van Rheenen, T.E., & Rossell, S.L. (2013a). Genetic and neurocognitive foundations of emotion abnormalities in bipolar disorder. *Cognitive Neuropsychiatry, 18*, 168–207. doi:10.1080/13546805.2012.690938
- Van Rheenen, T.E., & Rossell, S.L. (2013b). Is the non-verbal behavioural emotion-processing profile of bipolar disorder impaired? A critical review. *Acta Psychiatrica Scandinavica, 128*, 163–178. doi:10.1111/acps.12125
- Vaskinn, A., Sundet, K., Friis, S., Simonsen, C., Birkenæs, A.B., Engh, J.A., ... Andreassen, O.A. (2007). The effect of gender on emotion perception in schizophrenia and bipolar disorder. *Acta Psychiatrica Scandinavica, 116*, 263–270. doi:10.1111/j.1600-0447.2007.00991.x
- Vederman, A.C., Weisenbach, S.L., Rapport, L.J., Leon, H.M., Haase, B.D., Franti, L.M., ... McInnis, M.G. (2012). Modality-specific alterations in the perception of emotional stimuli in bipolar disorder compared to healthy controls and major depressive disorder. *Cortex, 48*, 1027–1034.
- Venn, H.R., Gray, J.M., Montagne, B., Murray, L.K., Michael Burt, D., Frigerio, E., ... Young, A.H. (2004). Perception of facial expressions of emotion in bipolar disorder. *Bipolar Disorders, 6*, 286–293. doi:10.1111/j.1399-5618.2004.00121.x
- Walker, E., McGuire, M., & Bettes, B. (1984). Recognition and identification of facial stimuli by schizophrenics and patients with affective disorders. *British Journal of Clinical Psychology, 23*, 37–44. doi:10.1111/j.2044-8260.1984.tb00624.x
- Young, R., Biggs, J., Ziegler, V., & Meyer, D. (1978). A rating scale for mania: Reliability, validity and sensitivity. *The British Journal of Psychiatry, 133*, 429–435. doi:10.1192/bjp.133.5.429