
Peer reviewed version

License (if available):
CC BY-NC-ND

Link to published version (if available):

Link to publication record in Explore Bristol Research
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via Elsevier at https://doi.org/10.1016/j.jad.2020.12.162. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research
General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/
Interpretation Bias Training for Bipolar Disorder:

A Randomized Controlled Trial
Abstract

Background

Bipolar disorder (BD) is associated with emotion interpretation biases that can exacerbate depressed mood. Interpretation bias training (IBT) may help; according to the “virtuous cycle” hypothesis, interpreting others’ emotions as positive can lead to interactions that improve mood. Our goals were to determine whether IBT can shift emotion interpretation biases and demonstrate clinical benefits (lower depressed mood, improved social function) in people with BD.

Method

Young adults with BD were recruited for three sessions of computer-based IBT. Active IBT targets negative emotion bias by training judgments of ambiguous face emotions towards happy judgments. Participants were randomized to active or sham IBT. Participants reported on mood and functioning at baseline, intervention end (week two), and week 10.

Results

Fifty participants (average age 22, 72% female) enrolled, 38 completed the week 10 follow-up. IBT shifted emotion interpretations (Hedges g=1.63). There was a group-by-time effect (-13.88, p<.0001) on self-reported depression; the IBT group had a larger decrease in depressed mood. The IBT group also had a larger increase in perceived familial support (3.88, p<.0001). Baseline learning rate (i.e., how quickly emotion judgments were updated) was associated with reduced clinician- (B=-54.70, p<0.001) and self-reported depression (-58.20, p=0.009).

Conclusion

Our results converge with prior work demonstrating that IBT may reduce depressed mood. Additionally, our results provide support for role of operant conditioning in the treatment of depression. People with BD spend more time depressed than manic; IBT, an easily disseminated intervention, could augment traditional forms of treatment without significant expense or side effects.

Key words

Bipolar disorder, social cognition, depression, social support, operant learning
Interpretation Bias Training for Bipolar Disorder: 
Building better mood through improved social cognition

Bipolar disorder (BD) is a serious mental illness, associated with significant burden to both individuals and society, and ranks as one of the leading causes of disability worldwide (Gore et al., 2011). Inability to regulate emotions is implicated as a key mechanism underlying the emotional lability of BD (Green, Cahill, & Malhi, 2007). People with BD show deficits in all stages of emotional processing, from detecting and interpreting emotion, to generating and regulating a response (Dickstein & Leibenluft, 2006).

Interacting with other people often cues emotional responses, and faces are one of the most salient categories of emotional stimuli (Dickstein & Leibenluft, 2006). In both adult and pediatric samples, BD is consistently associated with poor accuracy on face emotion recognition tasks (Schaefer et al., 2010; Schenkel, Pavuluri, Herbener, Harral, & Sweeney, 2007; Seymour et al., 2013; Van Rheenan & Rossell, 2014; Vederman et al., 2012). The results from some studies suggest that the nature of the misinterpretation may be related to mood state (i.e., a person who is depressed may interpret a neutral face as sad; Dickstein & Leibenluft, 2006; Lembke, 2002; Rich et al., 2005). Impaired attention may also play a role in the face emotion processing deficits observed in people with BD (Getz, Shear, Strakowski, 2003; Murphy et al., 1999). Misinterpreting others’ facial expressions can have a negative influence on social interactions, and BD is associated with significant social impairment across mood states (MacQueen, Young, & Joffe, 2001). For example, misinterpreting someone’s expression as angry, could lead to negative emotional responses from the person with BD, which would likely be off-putting or upsetting to the other person (Green et al., 2007). Interpretation biases may also exacerbate and perpetuate disordered mood states; when one is feeling low, seeing negative emotions in others could maintain a negative mood state (Lahera et al., 2012).
Emotion processing deficits are one challenge many individuals with BD face when trying to maintain relationships (Johnson, 1999). Mood symptoms also play a role; when manic, people with BD will have a tendency to engage in inappropriate behaviors, and when depressed, they may withdraw from other people (Michalak, Yatham, Kolesar, & Lam, 2006). Erratic behavior and not knowing what to expect from the person with BD are also cited as reasons that people with BD often struggle in their relationships (Michalak et al., 2006; Mortensen et al., 2015).

People with BD who manage to build and maintain strong relationships benefit; social support is associated with better outcomes (longer remission, fewer relapses; Cohen, Hammen, Henry, & Daley, 2004; Johnson, 1999), and most of the evidence-based therapies for both youth and adults with BD include interventions to capitalize on existing social support to help patients maintain stable mood. However, social deficits, including emotion perception errors and subsequent inappropriate responses, interfere with the ability of people with BD to benefit from most major interventions in which communication and interpersonal relationships are key (i.e., psychotherapy, recovery programs, illness management and recovery; Hasson-Ohayon, 2009).

Previous research has demonstrated that it is possible to shift emotion perception biases, and that doing so can have a positive effect on mood (Penton-Voak, 2012; Penton-Voak et al., 2013). In studies of adults with depressed mood, Penton-Voak et al. (2012; 2013) have demonstrated that by training participants to shift their perceptions facial expressions to be more positive, participants experience an increase in positive affect. A similar paradigm has been used successfully to train youth with aggressive behavior (Penton-Voak et al., 2013) or high irritability (Penton-Voak et al., 2013; Stoddard et al., 2016) to see more faces as happy, rather than angry, which in turn, reduced expressions of anger.

The use of cognitive remediation paradigms to reduce deficits of social cognition have been extensively tested in individuals with schizophrenia (Grant, Lawrence, Preti, Wykes, &
Cella, 2017; Horan & Green, 2019; Kurtz, Gagen, Rocha, Machado, & Penn, 2016) or autism spectrum disorders (Rice, Wall, Fogel, & Shic, 2015). However, a frequently noted limitation of these interventions is that the laboratory-based improvements do not translate into functional gains for patients. People with BD often have social deficits, but fewer social cognition domains tend to be affected than in people with schizophrenia or autism spectrum disorders (Addington & Addington, 1998; Couture et al., 2010; Lee et al., 2013). Additionally, people with BD typically desire social affiliation, which may not be true of people with autism spectrum disorders or schizophrenia (Blanchard, Park, Catalano, & Bennett, 2015; Chevallier, Kohls, Troiani, Brodkin, & Schultz, 2012). Consequently, remediating emotion perception deficits people with BD could have greater functional benefits than have been found in other populations. Previous studies have provided preliminary evidence for the use of social cognition interventions in people with BD, but have not evaluated secondary gains (e.g., improved mood) or whether benefits are sustained over time (Lahera et al., 2013).

In the present study our goal was to determine whether interpretation bias training (IBT; Penton-Voak, 2012; Penton-Voak et al., 2013), which has shown promise with depressed adults (cf. Penton-Voak et al., 2018) and irritable youth (Penton-Voak et al., 2013; Stoddard et al., 2016), can reduce the emotion perception biases in people with BD. Furthermore, we aimed to evaluate whether reducing the negative emotion interpretation bias in people with BD would lead to positive clinical changes. We hypothesized that seeing others’ emotions as happier would lead to improved social functioning and, consequently, reduce symptoms of depression as part of a “virtuous cycle” (Penton-Voak, 2012). We chose to train participants to see more faces as happy because people with bipolar disorder tend to spend three times as many days depressed as manic (Kupka et al., 2007), and the depressed mood episodes cause greater impairment (Michalak, Murray, Young, & Lam, 2008; Post, 2005). Directly targeting depressed mood has greater potential to improve functional outcomes in this population.
Method

Participants

Young adults (aged 16-25) with bipolar spectrum disorders were eligible. Participants with cognitive impairment (as assessed by a series of questions about their progress in mainstream school and whether they receive(d) special services in school), current psychosis, current suicidal intent, or current substance use disorder were excluded. All diagnoses were assessed using the Structured Clinical Interview for DSM-IV (SCID; First, Spitzer, Gibbon, & Williams, 1995).

Procedure

All procedures were approved by the Institutional Review Board. Participants were recruited through a medical center in the northeastern United States and through online advertisements. All potential participants (and/or a caregiver for those under 18), who expressed interest in the study, were interviewed on the phone to determine initial eligibility; those who met criteria were then invited for a comprehensive evaluation as part of their baseline appointment. Participants were randomized to either the active or sham IBT when their appointment was scheduled.

Following the informed consent/assent process, participants were interviewed by a doctoral student in psychology to assess clinical and treatment history, and to determine diagnostic eligibility, using the SCID. Additionally, current symptoms were assessed using the Young Mania Rating Scale (YMRS; Young, Biggs, Ziegler, & Meyer, 1978) and the Beck Depression Inventory, Second Edition (BDI-II; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961). Each interview was reviewed by the principle investigator, a licensed clinical psychologist, to ensure eligibility. Participants who met diagnostic and study eligibility requirements were also interviewed about their treatment history, and completed self-report measures assessing manic and depressive symptoms (General Behavior Inventory; Depue, Krauss, Spoont, & Arbisi, 1989), and perceived social support (Perceived Social Support from
Family and Friends; Lyons, Perrotta, & Hancher-Kvam, 1988). Interviewers were blinded to condition.

Following completion of the interview and self-report measures, participants completed the first session of IBT. Participants were scheduled to return for a second session of IBT within one week and for a third session of IBT and clinical assessment within two weeks. The two-week clinical assessment consisted of interviewing the participant with the YMRS and BDI-II, asking about any changes in treatment, and having the participant complete self-report measures of manic and depressive symptoms and perceived social support from friends and family. At this third appointment, participants were scheduled for their follow-up appointment, approximately eight weeks later. At the follow-up appointment, participants were again interviewed about their current symptoms and treatment, completed the same set of self-report measures, and repeated the IBT protocol.

Measures. Structured clinical interview for DSM-IV (First et al., 1995). The SCID is a semi-structured clinical interview designed to aid in the determination of Axis I disorders based on DSM-IV-TR criteria. It has excellent reliability and validity and is used widely for both research and clinical applications. All interviews were conducted by doctoral students in clinical psychology or by the principle investigator, a licensed clinical psychologist. The principle investigator reviewed all interviews to ensure reliability of the diagnoses.

Young Mania Rating Scale (YMRS; Young et al., 1978) was completed at the baseline appointment, the third appointment (end of intervention), and at the eight-week follow-up. Cronbach’s alpha was 0.73 at baseline.

Beck Depression Inventory II (BDI II; Beck et al., 1961) was administered as an interview at the baseline, third, and eight-week follow-up appointments. Cronbach’s alpha was 0.91 at baseline.
The General Behavior Inventory (GBI; Depue et al., 1989) participants completed the GBI at the baseline, third, and follow-up appointments. Cronbach’s alpha for the depression scale was 0.97 and for the hypomanic/biphasic scale was 0.91 at baseline.

Perceived Social Support from Family and Friends measure (Lyons et al., 1988; Procidano & Heller, 1983) was administered at the baseline, third, and follow-up appointments. Cronbach’s alpha for the family scale was 0.92, and for the friend scale was 0.46.

The Interpretation Bias Training (IBT; Penton-Voak, 2012) consists of two phases. In the first phase, participants view 15 images of a male face, which represent a morphed sequence from very sad to very happy. Each face is repeated three times for a total of 45 faces. When each face is presented, the participant must press a key to indicate whether s/he believes the face is happy or sad. Following the emotion choice, a static tile is shown as a mask, followed by another face. See Figure 1. The first phase allows for the determination of a participant’s balance point and his/her indifference point. The goal of the intervention is to train participants to see more faces as happy, not to make them more accurate per se. The hypothesis is that seeing others in a positive mood will induce a virtuous cycle, not that simply being more accurate is more desirable. Furthermore, there is no ground truth with these stimuli, as images in the center of the morph continuum are ambiguous.

Both statistics represent the same behavioral phenomenon, the point along the morph continuum where judgments change from predominantly sad to happy. However, they are estimated differently; balance point is calculated in-task as the proportion of sad to all judgments in the training block and then multiplied by the number of morphs in the continuum (15 in this case). This simple arithmetic estimate is based on the observation that most individuals shift their judgments abruptly along the morph continuum and it can be estimated easily within the task program. The indifference point is measured by fitting a four-point logistic curve to judgments. It closely adheres to choice theory (Luce, 1959), accommodating individuals who are less abrupt in
their transition, and accounting for asymmetric biases in judgments for morphs at the extremes of the continuum (Pollak & Kistler, 2002). However, it requires a nonlinear fit. The balance and indifference points are typically highly correlated (Deveney et al., 2019; Stoddard et al., 2016).

In phase two of the IBT task, the participant receives feedback about his/her emotion choices. The feedback is based on his/her balance point. For participants in the active IBT condition, the feedback aims to shift the balance point by two morphs, so that the participant interprets two more of the morphed faces as happy than s/he did during the phase one. For example, with 15 faces on the morph sequence, an individual might have a baseline balance point at face 9 – they judge faces 1-9 as sad and 10-15 as happy. In phase two, in order to be “right,” the participant must label faces 8-15 as happy (the original 10-15, plus two more). If any face numbered 8-15 is labeled as sad, the participant will get the feedback “Incorrect! That face was happy.” When faces are labeled correctly, based on the shifted balance point, s/he is told, “Correct! That face was sad.” For participants in the sham condition, the feedback reinforces their initial balance point and tells them they are correct when they respond in a way that is consistent with the first phase of the trial. There are six rounds of feedback in phase two, for a total of 225 faces. Following the training phase there is another assessment block of 45 trials identical to the first with no feedback to assess post-training balance point.

**Analytic plan.** IBT contains an internal quality check of accuracy (judgments of the four most extreme morphs corresponding to their overt expressions). If accuracy falls below 70%, it is likely the participant is poorly engaging (Stoddard et al., 2016). We used this externally developed performance criterion for subject-level inclusion.

We hypothesized that the main benefit of IBT would be reduced symptoms of depression due to the “virtuous cycle” through which improved social interactions result in more positive mood. Therefore, our main outcomes were defined as self- and clinician-rated symptoms of depression. Outcomes were evaluated using mixed models with maximum likelihood estimation;
we evaluated the efficacy of the intervention by testing group (IBT or sham), time, and group*time as predictors. Additionally, we evaluated the effect of IBT on perceived social support (friends and family), and tested mediation with bootstrapping to get unbiased estimates of the direct and indirect effects of IBT and perceived social support on depressed mood (Page-Gould, 2016). Finally, we tested whether there was any effect of IBT on self- or clinician-rated manic symptoms as a sensitivity analysis. We included sex as a covariate, given sex differences in emotion processing (Craig et al., 2009; Montagne, Kessels, Frigerio, de Haan, & Perrett, 2005; Thayer & Johnsen, 2000). All participants were included in the mixed models, regardless of whether they completed the intervention.

As described above, although IBT requires balance point to estimate initial bias, the indifference point is a more valid measure of the point at which participants distinguish between happy and sad faces, with respect to choice theory and precision. Recently a computational model of associative learning (Kruschke, 1992) has been developed for IBT allowing us to measure the speed at which feedback is incorporated into new responses (learning rate; https://github.com/joelStod/FaceEmo). Participants’ ability to quickly update their emotion interpretations based on the feedback provided is related to their clinical outcomes (i.e., people who learn more quickly are more likely to improve symptomatically). Interestingly, this effect is not limited to the intervention group; in the sham training, participants still receive feedback, which may encourage them to pay closer attention to the face stimuli thereby improving their attention to nuanced features of facial expressions and facilitating the ability to better discriminate between emotions. Therefore, a secondary aim of the study was to evaluate how learning rate in the initial session related to outcomes.

Results

Ninety participants were scheduled for a baseline appointment. Of these, 59 attended a baseline appointment, 55 were eligible to participate. An additional five people were eliminated from the
analyses because their baseline assessment of emotion recognition was less than 70% accurate. Therefore there were 50 people in the baseline group. Three participants withdrew following session one, due to transportation or time constraints, two were no shows, leaving 45 participants with valid data at session two. Forty participants had valid data at the third session, completing the intervention. Thirty-eight participants had valid data at the eight-week follow-up. Figure 2.

The average age of the sample was 22 years (SD=2.4), 72% (n=36) were female. The majority (n=35) had a diagnosis of bipolar I, nine had bipolar II, five had other specified bipolar disorder and one participant had cyclothymic disorder. Nineteen participants were currently depressed, 3 were hypomanic, and 5 were experiencing mixed symptoms. Half of the participants reported their race as “other,” 12 were White, 11 were Black, two were Asian, and 32 reported their ethnicity as Hispanic. Almost a third (n=14) of the participants were not in treatment; 29 reported taking medication and 26 were in therapy (20 were receiving therapy and medication). None of these clinical or demographic variables varied between the IBT and sham groups. Participant scores on the measures of mood symptoms and social support were also equivalent across groups, with the exception of self-reported depression, which was higher in the IBT group (t=2.75, p=.008, Hedges g=0.79). Table 1.

The two groups did not differ in average balance point or indifference point at baseline, but – as expected – the IBT group had significantly larger shifts in their balance point (t=5.84, p<.0001, Hedges g=1.85) and indifference point (t=5.17, p<.0001, Hedges g=1.63), indicating that the intervention was successful at prompting more happy judgments of ambiguous faces. Of note, the balance point and indifference point were highly correlated (r=0.96, p<.0001), but the balance point was less than the indifference point (mean difference = 0.42, t=2.68, p=.007).1

1 One source for a systematic difference is that the balance point is estimated by a simple proportion and then scaled to the morph continuum by multiplying it by the total number of morph intervals, here 15. However, morphs are linearly spaced and indicated by integers. So, a simple correction for the starting value, 1, eliminates much of this difference. For example, calculating balance point here as
There was weak evidence that indifference point was correlated with clinician-rated depression at baseline ($r = -0.27, p = .063$), but not for self-reported depression ($r = -0.10, p = .486$), mania (self-report [$r = -0.20, p = .178$], or clinician-rated [$r = -0.14, p = .334$]) or perceived social support (friend ($r = 0.00, p = .999$), or family($r = 0.16, p = .276$)). At the eight-week follow-up, change in indifference point was not related to change in any of the outcomes.

Existing treatment, including psychotherapy and psychotropic medication, were not associated with change in any of the outcome measures, with the exception of perceived familial support; individuals in psychotherapy reported less change in perceived familial support than those who were not in therapy ($t=2.18, p=0.036$).

Next, we evaluated the association between the intervention and primary (depressive symptoms) and secondary (perceived social support) outcomes using mixed models to allow for random intercepts among participants and missing data across time points, see Table 2. The model predicting self-reported depression scores showed a significant group*session interaction, with the IBT group experiencing significantly greater improvements by the eight-week follow-up, see Figure 3. The model predicting perceived familial support had a significant group*session interaction, with the IBT group showing improved familial support at the eight-week follow-up session. None of the hypothesized predictors were significantly associated with perceived friend support. See Figure 4.

Changes in self-reported depression and perceived familial support were weakly associated ($r=-0.27, p=0.095$). However, the model testing familial support as a mediator between IBT and self-reported depression outcomes indicated that although IBT was associated with familial support ($B=3.37, p=0.036$) and self-reported depression ($B=-12.13, p=0.017$), the path between familial support and self-reported depression was not significant ($B=-0.49, p=0.305$). In

\[
\frac{n_{\text{sad}}}{n_{\text{total}}} (15 - 1) + 1, \text{ where } n \text{ refers to the number of sad or all judgments. This correction eliminates the mean difference (mean difference = 0.06, } t=0.37, p=0.71).\]
contrast, changes in clinician-rated depression and perceived familial support were not correlated ($r=0.01$, $p=0.942$), and in multilevel path analysis, IBT was associated with improved support ($B=3.15$, $p=0.08$), but IBT was not associated with depression scores ($B=0.11$, $p=0.975$) and familial support was not associated with depression ($B=0.02$, $p=0.937$).

In the model predicting self-reported manic symptoms, there was a trend-level difference for group, with the IBT group showing lower manic symptoms at the eight-week follow-up session. There was a main effect of IBT on clinician-rated manic symptoms, with the IBT group showing lower symptoms, but the group*session interaction showed higher manic symptoms in the IBT group at session three.

To examine ability to learn, we conducted mixed effects models consistent with the main analysis, predicting outcomes by the fully interactive effects of baseline learning, IBT condition, and session with gender as a covariate. A three way interaction between baseline learning, IBT condition, and 8-week follow up was associated with clinician-rated depression ($B=-54.70 [-75.97\ –\ 33.44], p<0.001$), self-reported depression ($B=-58.20 [-102.27\ –\ -14.13], p=0.009$), clinician rated mania ($B=19.27 [8.74\ –\ 29.79], p<0.001$), perceived family support ($B= 18.30 [5.81\ –\ 30.79], p=0.004$). No three-way interaction was associated with self-reported mania or perceived peer support. In post hoc analyses contrasting condition (active – sham) and clinical evaluation (last – first), adjusting for baseline learning, clinician- and self-rated depression decreased, perceived family support and clinician-rated mania symptoms increased ($ps<.01$). Across sessions, for those in the active IBT condition, increased learning rate was associated with decreased depression ($r=-0.54, p=0.008$).

**Discussion**

The goal of the present study was to evaluate the efficacy of a computer-based interpretation bias training intervention to modify negative bias in the interpretation of others’ emotions among young adults with bipolar disorder. This intervention has been used successfully in other
IBT FOR BIPOLAR DISORDER

populations to promote positive (i.e., happy) emotion interpretations (Adams, Penton-Voak, Harmer, Holmes, & Munafò, 2013; Penton-Voak, 2012; Penton-Voak et al., 2013). Additionally, these studies have found evidence of a “virtuous cycle” whereby positive interpretations of others’ emotions lead to more successful social interactions, which in turn, improves mental health. We found that IBT successfully modified participants’ emotion interpretations with a large effect size maintained over a two-month follow-up. However, effects on participants’ mood and social functioning were mixed.

The social cognition deficits among people with BD have been well described (Martino, Strejilevich, Fassi, Marengo, & Igoa, 2011; Samamé, 2013; Samamé, Martino, & Strejilevich, 2012), and these deficits are associated with poor social functioning (Purcell, Phillips, & Gruber; Van Rheenen & Rossell, 2014) and depressed mood (Lahera, 2012). Additionally, recent work suggests that deficits in recognizing others’ facial emotions may be especially pronounced among people with BD and that accurate interpretation of others’ emotions is negatively correlated with mood symptoms and positively correlated with quality of life and social functioning (Lee & Van Meter, 2020). The fact that IBT successfully modified negative bias in emotion judgments and that this change was maintained over a two-month follow-up is significant.

People with BD will spend as many as three times more days depressed than manic or hypomanic (Kupka et al., 2007). This fact was a key factor in the decision to target depressed, rather than irritable/angry, mood in this trial. The IBT group did experience significant improvements in their self-reported depressed mood over time, consistent with study goals. However, changes in clinician-rated depression scores did not differ between groups. It’s possible this was related to reluctance to endorse depressive symptoms in an interview format (Hunt, Auriemma, & Cashaw, 2003; Kaplan et al., 1994; Newman et al., 2002). Additionally, the self-report measure has more items than the clinician rating scale, which may better capture the experience of depressed mood. It is also important to note that the IBT group had higher rated
depression at baseline and therefore some improvement could be attributed to regression to the mean.

Based on the “virtuous cycle” theory, we hypothesized that participants in the active IBT group would experience improvements in their social functioning. This hypothesis was partly supported. Participants in the IBT group reported improvements in their family relationships, but not with peers, and although family support and depression were weakly associated over time, family support did not mediate the association between the intervention and change in depression. It may be that, because the hypothesized mechanism of change works through improved social cognition and reduced mood symptoms, not enough time had passed (eight weeks) for the full effect of the intervention to take hold. Previous studies of IBT have found that the benefits may increase over time, as the emotion perception skills are practiced and reinforced (Adams et al., 2013; Penton-Voak et al., 2013). Consequently, a longer follow-up period may have revealed a greater impact on both mood and social functioning. It is also possible that effects were seen in familial relationships, but not peer, because there is greater motivation with peers to censor behavior and act appropriately, whereas interactions with family members may more affected by misinterpretations and low mood, creating a greater opportunity for improvement.

We hypothesized that participants’ ability to learn – to update their judgments of others’ faces based on feedback – might be an important determinant of outcome. This hypothesis was largely supported; with more efficient learning modulating the association between improvement in clinician-and self-rated depression and improvement in perceived familial support in the active training group. This is consistent with prior work that the ability to learn during operant conditioning is associated with positive outcomes in depression (Pizzagalli, Goetz, Ostacher, Iosifescu, & Perlis, 2008; Vrieze et al., 2013). It may be that it is the ability to respond to operant learning or to notice subtle emotion cues in faces that is more important, rather than simply judging more faces as happy. An important next step will be to evaluate whether IBT training
generalizes to other emotions – i.e., does social cognition improve broadly, not just for sad and happy faces.

Compared to other treatments available to people with BD, IBT requires fewer resources than typical psychotherapy, can be reliably delivered, including online (Stoddard et al., 2016), and does not have any of the harmful side effects associated with some pharmacological interventions. Although effective psychosocial treatments for BD exist (Fristad & MacPherson, 2014; Schottle, Huber, Bock, & Meyer, 2011) dissemination has been limited, and the time and resource commitment is high. Additionally, young adults, the age group targeted in the present study, tend to be among the most difficult to engage in traditional mental health services (Heflinger & Hinshaw, 2010; McGorry, Bates, & Birchwood, 2013). Developing alternative methods of effectively intervening in this population is important (Kazdin & Blase, 2011). Internet-based therapies are gaining popularity (Andersson, Carlbring, Ljótsson, & Hedman, 2013), and may be more appealing to young adults (Baruch, Vrouva, & Fearon, 2009). Although the present study was conducted in a lab, the intervention has been adapted for web-based administration, which facilitates dissemination and would enable the administration of the intervention at patients’ homes, pending evidence of efficacy from in-home trials. Although IBT is not conceived as a replacement for psychotherapy or other treatments for BD, it does successfully target a consequential deficit and has the potential to improve outcomes important to patients, including social functioning and quality of life.

Our results indicate that it may be possible to train people with BD to correct their emotion perception biases in order to increase positive mood and promote mood stability, however the lasting effects of this intervention are unknown. In a study using the same IBT paradigm, neuroimaging results showed increases in whole brain, medial prefrontal cortex (mPFC), and bilateral amygdala activation in response to happy faces six weeks following the intervention (Penton-Voak et al., 2018), and in a similar study of angry-to-happy bias
modification, changes in activation in the lateral orbitofrontal cortex were found post-intervention (Stoddard et al., 2016). These findings are similar to results from studies of antidepressant medication, which is encouraging (Warren, Pringle, & Harmer, 2015). Research on similar interventions in other clinical populations show similar findings; structural and functional changes were measured following cognitive remediation in the brains of people with schizophrenia (Penadés et al., 2013). Importantly, these effects can be long-lasting; Eack et al. (2010) found higher grey matter volume in patients who received cognitive remediation two years after the intervention. However, longer follow-up is necessary to better evaluate the “virtuous cycle” hypothesis and whether effects increase and persist over time.

This study is the first to evaluate IBT in young people with bipolar disorder. Although improvements in self-reported depression and familial support are encouraging, meaningful differences in clinician-rated scales were not observed, thus results must be interpreted with caution. Additionally, before IBT should be implemented with patients, additional trials are necessary to address the limitations of our investigation, which may have impacted outcomes. Participants were not selected on emotion interpretation bias, and some were 100% accurate to start. However, these individuals could still benefit from IBT, as the goal is to increase the number of other people who are seen as being happy, not to improve accuracy. Related, participants were not recruited based on current symptoms or social functioning. Consequently, there may have been a ceiling on the degree to which these domains could improve. Additionally, participants were functioning at a level sufficient to attend four sessions, which may have biased the sample with higher functioning individuals. In the future, it will be important to recruit a larger sample of participants with current mood symptoms and to evaluate the social cognition outcomes more closely, in order to better understand the mechanism of change.

Social cognition deficits are present early in the course of bipolar disorder (Pavuluri, O'Connor, Harral, & Sweeney, 2007; Rich et al., 2010; Rich et al., 2008), and social functioning
deficits in this population contribute to mood problems (Goldstein et al., 2009; Keenan-Miller & Miklowitz, 2011; Keenan-Miller, Peris, Axelson, Kowatch, & Miklowitz, 2012). A low burden, augmentive intervention to improve social cognition near the time of initial mood episode onset could have a large impact by enabling youth to achieve adequate social functioning and to benefit from the support it provides, including better functioning and longer periods of wellness (Eidelman, Gershon, Kaplan, McGlinchey, & Harvey, 2012; Goldstein, Miklowitz, & Mullen, 2006; Johnson, 1999). Although research tends to focus on the etiology and direct treatment of mood symptoms, there may be other paths through which functioning and quality of life can be improved. Our results converge with prior work demonstrating IBT may reduce depressed mood (Penton-Voak, 2012). People with BD spend more time depressed than manic, and depressed mood tends to be more difficult to treat (Baldessarini, Vieta, Calabrese, Tohen, & Bowden, 2010; Van Meter, Henry, & West, 2013); IBT, a computer-based, easily disseminated intervention, has promise as an augmentive therapy for people with BD without significant expense or side effects.
Table 1.

Demographic and clinical characteristics of the sample

<table>
<thead>
<tr>
<th>Interpretation Bias Training</th>
<th>Sham</th>
</tr>
</thead>
<tbody>
<tr>
<td>n (%)</td>
<td></td>
</tr>
<tr>
<td>Demographic information</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>21 (68)</td>
</tr>
<tr>
<td>Asian</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Black/African American</td>
<td>7 (23)</td>
</tr>
<tr>
<td>White/Caucasian</td>
<td>6 (19)</td>
</tr>
<tr>
<td>Other race/multiracial</td>
<td>17 (55)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>22 (71)</td>
</tr>
<tr>
<td>Clinical information</td>
<td></td>
</tr>
<tr>
<td>Bipolar I</td>
<td>19 (61)</td>
</tr>
<tr>
<td>Bipolar II</td>
<td>8 (26)</td>
</tr>
<tr>
<td>Other specified bipolar disorder</td>
<td>3 (10)</td>
</tr>
<tr>
<td>Cyclothymic disorder</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Current episode</td>
<td></td>
</tr>
<tr>
<td>Depressed</td>
<td>14 (45)</td>
</tr>
<tr>
<td>Hypomanic</td>
<td>2 (6)</td>
</tr>
<tr>
<td>Mixed</td>
<td>3 (10)</td>
</tr>
<tr>
<td>Taking medication</td>
<td>19 (61)</td>
</tr>
<tr>
<td>Baseline clinical ratings</td>
<td>Mean (SD)</td>
</tr>
<tr>
<td>Clinician-rated depression</td>
<td>20.17 (13.0)</td>
</tr>
<tr>
<td>Self-reported depression</td>
<td>116.73 (26.1)</td>
</tr>
<tr>
<td>Clinician-rated Mania</td>
<td>14.93 (3.2)</td>
</tr>
<tr>
<td>Self-reported Mania</td>
<td>65.17 (15.4)</td>
</tr>
<tr>
<td>Social Support - Family</td>
<td>8.61 (6.2)</td>
</tr>
<tr>
<td>Social Support - Friends</td>
<td>13.13 (3.9)</td>
</tr>
<tr>
<td>Baseline task performance</td>
<td>Mean (SD)</td>
</tr>
<tr>
<td>Balance Point</td>
<td>6.67 (11.6)</td>
</tr>
<tr>
<td>Indifference Point</td>
<td>7.12 (1.7)</td>
</tr>
<tr>
<td>Learning rate</td>
<td>0.92 (0.8)</td>
</tr>
</tbody>
</table>
Table 2. Mixed model outcomes

<table>
<thead>
<tr>
<th></th>
<th>BDI</th>
<th>GBI Depression</th>
<th>YMRS</th>
<th>GBI Mania</th>
<th>Perceived Social Support</th>
<th>Perceived Social Support</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>B (95% CI)</td>
<td>p-value</td>
<td>B (95% CI)</td>
<td>p-value</td>
<td>B (95% CI)</td>
</tr>
<tr>
<td>Intercept</td>
<td>15.14</td>
<td>(10.1-20.1)</td>
<td>0.000</td>
<td>97.58</td>
<td>(85.1-110.0)</td>
<td>0.000</td>
</tr>
<tr>
<td>IBT</td>
<td>5.27</td>
<td>(0.8-11.4)</td>
<td>0.102</td>
<td>20.65</td>
<td>(5.5-35.8)</td>
<td>0.012</td>
</tr>
<tr>
<td>Gender</td>
<td>-2.52</td>
<td>(-9.2-4.2)</td>
<td>0.471</td>
<td>-5.80</td>
<td>(-22.7-11.1)</td>
<td>0.510</td>
</tr>
<tr>
<td>Session 2</td>
<td>-4.35</td>
<td>(-6.9-1.8)</td>
<td>0.001</td>
<td>-3.09</td>
<td>(-8.1-1.9)</td>
<td>0.233</td>
</tr>
<tr>
<td>Session 3</td>
<td>-3.89</td>
<td>(-6.4-1.4)</td>
<td>0.003</td>
<td>-3.18</td>
<td>(-1.9-8.3)</td>
<td>0.231</td>
</tr>
<tr>
<td>Session 2*IBT</td>
<td>3.26</td>
<td>(-0.0-6.5)</td>
<td>0.056</td>
<td>-1.77</td>
<td>(-8.2-4.6)</td>
<td>0.593</td>
</tr>
<tr>
<td>Session 3*IBT</td>
<td>-0.16</td>
<td>(-3.4-3.1)</td>
<td>0.923</td>
<td>-13.88</td>
<td>(-20.4-7.3)</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Random Effects

|                      | σ²   | 24.45 | 100.02 | 6.35 | 35.69 | 7.71 | 6.25 |
|                      | τ00 ID | 100.98 | 652.67 | 7.93 | 210.63 | 29.05 | 14.59 |
| Intraclass correlation | 0.81 | 0.87 | 0.56 | 0.86 | 0.79 | 0.7 |

Note: BDI – Beck Depression Inventory-II, GBI – General Behavior Inventory, YMRS – Young Mania Rating Scale, IBT – Interpretation Bias Training
Figure 1.

*Face morph sequence from sad-to-happy*
Figure 2. CONSORT diagram
Figure 3. Change in self-reported depression symptoms

*** $p < .0005$
Figure 4. *Change in perceived support from family*

![Graph showing change in perceived support from family over sessions. The graph indicates a significant change with an asterisk Annotations: ***p < .0005***]
Acknowledgements

Thank you to Neha Agrawal, Michelle Chu, Yana Lechtman, Jun Min Koay, Patricia Lee, Melissa Roed, Miranda Rosenberg, and Joey Sergi for their help with data collection.
References


