Authors: *First, we would like to thank all reviewers for their favourable evaluations and helpful suggestions. Based on the reviewers' comments, major improvements include additional PPI analyses, the inclusion of behavioral data (reaction times) as well as the implementation of an anhedonia scale. Additionally, an analysis of covariance (ANCOVA) was conducted on NAcc responsiveness to investigate the potential effects of clinical variables on our results. This was complemented by separate correlation analyses between clinical variables and NAcc responsiveness to explore possible associations. Furthermore, to study the applicability of the present findings in discriminating UD and BD, a linear discriminant analysis was conducted.*

We are convinced that these additions have strengthened our manuscript substantially and we hope that it will now meet the high standards of Neuropsychopharmacology. Specific comments to all reviewer questions and suggestions can be found below.

Reviewer #1 :

Redlich et al compared UD and BD during reward processing. The authors found lower activity in the NAcc in UD and BD compared to HC and in BD compared to UD. They also found greater connectivity between NAcc and VTA when comparing UD with HC and a trend between BD and HC.

The authors used state of the art imaging methodology to address a gap in the literature, ie direct comparison between Unipolar depression and Bipolar depression. Only one study has previously compared UDvsBD during reward processing using fmri. Therefore, the current work is a major contribution for the field. Furthermore, they also explored connectivity between regions, which allows the authors to investigate neuro circuits implicated with reward processing.

I have only minor comments that could potentially enhance this already great paper.

1. the NAcc abbreviation should be explained when it first appears.

Authors: *We thank the reviewer for that hint. We now explain the abbreviation when it first appears in the introduction.*

2. The section questionnaire measures and rating scales is somewhat lost in the manuscript. The authors could explore the significant activity/connectivity with these clinical variables to see if it is driving the results. The childhood maltreatment measures could be excluded from the manuscript since it is not the main focus of the paper.

Authors: *Thank you for pointing out this lack of clarity. We incorporated the section "questionnaire measures and rating scales" into the section 2.1 ("Participants and Questionnaires"). The childhood trauma questionnaire (CTQ) was excluded from the manuscript. To detect if questionnaire measures and clinical variables were driving the group differences between BD and UD regarding NAcc responsiveness to reward stimuli, the peak contrast values of the 3 (group) x 2 (condition) interaction analysis of the bilateral NAcc and the significant cluster from the functional connectivity analysis were extracted for each participant and further analyzed with PASW Statistics 22 (IBM, Armonk, New York). An analysis of covariance (ANCOVA) was conducted on NAcc responsiveness with the factor group (UD, BD) and covariates BDI, HAMD, SHAPS-D, medication load index, number of depressive episodes, and time since onset of first depressive episode. Adding these clinical*

variables as covariates only slightly weakened the results regarding the condition × group interaction (NAcc right: P = 0.003, NAcc left P = 0.014).

3. Discussion: the connectivity findings and interpretations are interesting and could be further discussed. For example, another explanation is excessive prefrontal cortex regulation upon the NAcc during the anticipatory part of the task (that was not explored in the manuscript). You could include in the limitation paragraph this point since it can be explored using Dynamic Causing Modeling within SPM.

Authors: *The reviewer addresses an interesting interpretation: the excessive prefrontal cortex regulation upon the NAcc. Previous studies demonstrated a strong prefrontal cortex modulation upon the NAcc (Richard and Berridge, 2013), that could explain the reduced NAcc activity despite a potentially stronger neural projection from the VTA. We added this alternative explanation to our discussion.*

Since just one previous study investigated reward processing directly comparing UD and BD, we followed a more exploratory approach (whole-brain functional connectivity using functional connectivity/ PPI analyses) rather than a confirmatory approach like DCM (Friston *et al*, 2003)*, which is rather model dependent and hypothesis driven. However, we agree with the reviewer that DCM is definitely interesting and should be applied in future studies which we now address in our discussion section.*

4. Furthermore, the authors mention that differences in reward activity could be due to a greater impairment in the mesolimbic system in BD. However, it could be also due abnormal regulation, that are not necessary present during reward outcome. See similar regulation abnormality in Almeida et al 2009 "Abnormal amygdala-prefrontal effective connectivity to happy faces differentiates bipolar from major depression".

Authors: *We thank the reviewer for providing these insightful thoughts. Almeida et al. (2009) found asymmetrical differences in effective connectivity between the OMPFC and amygdala in patients with BD and UD. Similar regulation processing could play an important role in the course of reward processing and lead to the presented outcome. We added this alternative explanation to our discussion.*

5. In the figure, please indicate significant post hoc results with "lines and asterisks"

Authors: *We thank the reviewer for this suggestion. We indicated significant differences within the figure with lines and asterisks for a better understanding.*

Reviewer #2:

Differentiating between bipolar (BP) and unipolar (UP) depressed patients is a major clinical challenge. In light of hypothesized reward processing dysfunction across the mood disorder spectrum, the authors used a card guess paradigm in 33 UP, 33 BP, and 34 healthy controls to investigate putative differences between depressed individuals with UP vs. BP. Relative to both the UP and HC group, the BP group showed reduced activation in the nucleus accumbens in response to rewards (but not losses), as well as reduced activation in other reward-related regions, such as the caudate, putamen and insula. In addition, relative to HC,

the UP group also showed reduced NAc activation to rewards. Finally, relative to HC, the UP group showed higher functional connectivity between the NAc and VTA. Overall, the study was well conducted and the findings are interesting, albeit somewhat counterintuitive (see below). Strengths of the study include the study of sizable sample across two DSM disorders, well-matched groups, and the use of standard methods. In spite of these strengths, some issues diminish the impact of this interesting contribution:

1) Findings are somewhat counterintuitive (most blunted NAc responses in BP; increased functional NAc-VTA connectivity in MDD). Additional explanations are needed, especially in light of recent findings and reviews highlighting increased NAc responsiveness across mood states in BP (e.g., Nusslock et al., 2014, BRAT).

Authors: *We thank the reviewer to draw our attention to this review. We added this reference that now strengthened our manuscript by underlining the complexity of reward processing in bipolar disorder. It seems to be of great importance to separate the investigation of reward anticipation and reward outcome to receive less counterintuitive findings. We added these thoughts to our discussion on page 12.*

2) Did groups differ in smoking status? In light of increased rates of smoking in mood disorders (and in particular, in BP) and emerging findings of the effects of nicotine (and craving) on striatal functions, this information is of paramount importance.

Authors: *We agree with the reviewer that the effects of nicotine might have an impact on striatal structure and function. Unfortunately, smoking status was not obtained in the context of the current study. Thus we are not able to make any statements about smoking and state this as a limitation within the discussion on page 16.*

3) In schizophrenia, typical and atypical antipsychotics have been found to differentially affect striatal responsiveness in similar tasks. Owing to the fact that a significantly higher proportion of BP patients were treated with antipsychotics, there is the distinct possibility that the present findings are confounded by medication effects. These control analyses are needed in order to fully evaluate the findings.

Authors: *The reviewer raises the important point of medication. Since the majority of patients with bipolar disorder were treated with antipsychotics (n = 24) or mood stabilizer (n = 17) we agree with the reviewer, that the present findings could possibly be confounded by medication effects. To control for a possible influence of medication we added the medication load index score as described earlier (Redlich et al, 2014), reflecting dose and variety of different medication, as covariate to an analysis of covariance (ANCOVA) on NAcc responsiveness with factor group (UD, BD) and covariates (BDI, HAMDs, SHAPS-D, medication load index, number of depressive episodes, time since onset of first depressive episode). This analysis revealed that adding these clinical variables as covariates only slightly weaken the results regarding the condition × group interaction (NAcc right: P = 0.003, NAcc left P = 0.014) (page 11, also see Reviewer #1, Point 2). Additionally, we calculated a t-test (using the contrast = reward > control) comparing bipolar patients treated with antipsychotic drugs (n=24) with bipolar patients treated without antipsychotic drugs (n=8). This analysis yielded no significant differences in NAcc activity between patients treated with antipsychotic drugs versus those who were not treated with antipsychotic drugs. This is in line with studies reporting that psychotropic medication appears to have only a limited impact on fMRI results (Hafeman, Chang, Garrett, Sanders, & Phillips, 2012; Phillips, Travis, Kupfer, & Fagiolini, 2008).*

Nevertheless, we cannot completely rule out specific medication effects. In light of studies that demonstrate a reduction of reward-related activation by antipsychotics in regions such as the ventral striatum (Abler *et al*, 2007)*, it is possible that our findings are still confounded. We address that mood stabilizers, antipsychotics, and antidepressants might each have different neurobiological effects in the limitation section.*

4) At the bottom of page 4, the authors introduce the hypotheses, which pertain to "reward processing". They should be more specific with respect to the phase of reward processing expected to differentiate the groups.

Authors: *We thank the reviewer for pointing us to the ambiguous phrase. Accordingly, we have specified our hypotheses.*

5) It is unclear whether the authors performed simple functional connectivity analyses or PPI analyses. The latter would provide more detailed insights of putative connectivity differences.

Authors: *We thank the reviewer for pointing out this unclarity. We applied a general functional connectivity approach, independently from task conditions. However, we agree with the reviewer that a PPI analysis is interesting, particularly with respect to our research question. Therefore, we calculated a psychophysiological interaction analysis (PPI) examining connectivity as a function of condition to detect possible patterns of altered network connectivity related to the reward versus control condition as employed by Laeger et al.* (Friston *et al*, 1997; Laeger *et al*, 2014)*. Again, the left nucleus accumbens was defined as seed region, reflecting the physiological variable. The contrast vector for reward > control was entered as psychological variable. For each subject a PPI was generated reflecting the increase of NAcc connectivity in the reward condition compared to the control condition. The resulting contrast images were entered into a one-way ANOVA using group (UD vs. BD vs. HC) as between-subjects factor. However, no significant differences could be found even when using a more liberal threshold ($p < 0.001$, $k > 50$). This result may indicate that the reported alterations of the VTA-NAcc-pathway in both patient groups could be due to general alterations of connectivity and are not reflecting dynamic changes in connectivity based on reward condition. Reviewer #1 suggested an interesting alternative interpretation of the connectivity findings: excessive prefrontal cortex regulation upon the NAcc. Recent studies demonstrated a strong prefrontal cortex modulation upon the NAcc (Richard and Berridge, 2013), that could explain the reduced NAcc activity despite a possible higher innervation by the VTA (please see Reviewer #1, Point 3). In light of the PPI results, this higher innervation seems to be a general alteration in the mesolimbic system independent of reward conditions.*

6) Did the UP and BP group truly excluded individuals with past substance abuse? In light of the high comorbidity between BP and substance abuse, this exclusion is rather surprising.

Authors: *Comorbid life-time diagnoses of any substance-related disorders both abuse and dependence were exclusion criteria (with the exception of nicotine abuse). This has been confirmed by the structured clinical interview for DSM-IV (SCID-IV; Wittchen et al, 1997). However, the reviewer is very right pointing to the high co-morbidity of BD and substance related disorders which caused several otherwise eligible patients not to be included in our study.*

7) In light of the study focus, it is surprising that no anhedonia scale was administered. If such scale was given, the authors should report the findings.

Authors: *We thank the reviewer for this valuable suggestion and implemented this advice. Indeed, among the available data in this ongoing study, also the German version of the Snaith-Hamilton Pleasure Scale (SHAPS-D,* Franz *et al*, 1998) *was employed and the corresponding data are now presented in table 1. While there were no significant differences between the UD sample and the BD sample (P = .07), the HC expectedly showed less anhedonia and significantly differed from UD as well as BD patients (Ps < 0.001). Stronger NAcc-VTA connectivity was associated with SHAPS scores (r = 0.305, P = .002) with respect to the whole sample (including the HC). However, there were no significant correlations between SHAPS-D scores and activity/connectivity measures in the patient groups alone, neither for UD nor for BD (Ps > .21, see supplemental table 2). We thank the reviewer for the suggestion.*

8) It could be interesting to conduct discrimination analyses entering NAc BOLD responses and NAc-VTA functional connectivity measures to differentiate the UP and BP groups.

Authors: *The reviewer raises an interesting point. Thus, we performed ta a discriminant analysis to classify patients with UD and patients with BD based on the mean contrast values of the bilateral NAcc and the significant cluster from functional-connectivity analysis (VTA) using PASW statistics 22 (IBM, Armonk, New York). The discriminant analysis yielded to 66, 7 % accuracy rate (sensitivity, correctly classified patients with UD = 74.2 %; specificity, correctly classified patients with BD = 59.4 %; eigenvalue = 0.25; Wilks's $\lambda = 0.80$; P = 0.004). Although the accuracy rate of 66,7 % may not be regarded as substantial enough to justify a clinical application, this result supports our univariate findings and confirms the importance of this system to differentiate both disorders.*

Minor points:

1) Did the authors use the same task as Chase et al.? If so, please clarify, so that readers can compare findings.

Authors: *We apologize for the lack of clarity. Chase et al. used a paradigm that included both anticipation and feedback phase, whereas our paradigm consists only of a feedback phase. We have now clarified this in both introduction and methods.*

2) Page 8: "Welcome" should be "Wellcome"

Authors: *We have corrected this mistake.*

3) Page 13: The sentence starting with "Given the differences..." is speculative.

Authors: *We thank the reviewer for pointing that out. We have changed that sentence and softened the statement.*

4) In Table 1, specify whether the t-tests pertain to MDD vs. BP differences.

Authors: *The reviewer is right, the t-tests and X^2-tests in table 1 refer to the comparison between the MDD sample and the BP sample. We have specified this in the table.*

5) Did the number of episodes or the lifetime cumulation of MDE modulate the findings?

Authors: *Our analysis of covariance (on page 12-13) yielded no significant modulation of neural activity/connectivity by the number of depressive episodes or time since onset of first depressive episode (please see also Reviewer #2, Point 3; Reviewer #1, Point 2).*

6) Did groups differ in the duration of the current episode?

Authors: *We thank the reviewer for this comment regarding table 1. Both groups did not differ in the duration of current episode (P = 0.089; UD, M = 14.09, SD = 13.20; BD, M = 23.76, SD = 29.31), which is now corrected in table 1.*

7) Did groups differ in SES?

Authors: *Regarding total education time in years, which was obtained as one indicator of SES, the three groups did not differ significantly ($F_{(2, 93)}$ =1.11, P =.333) and seem to be very well matched (UD-Mean: 14.76; BD-Mean =14.18, HC-Mean =14.85). However, even if education is a good predictor of income and socio-economic status later in life, we cannot rule out specific associations of current income or occupation on the present findings.*

Reviewer #3:

The manuscript is well-written, logically organized, and adequately illustrated. Overall it's a solid paper with novel findings in the field of mood disorders thanks to the application of both functional MRI and functional connectivity measures. Two topics should however be addressed in the results and/or conclusions as I think they are essential to help move the field (in particular treatment/prevention oriented research) forwards.

1. Did the authors attempt to relate neural measurements to behavioural findings? I noticed that reaction times/accuracy levels are not reported in either tables/figures or the results section. If not can they discuss this topic in the conclusions? For instance, cognitive impairment associated with reward processing deficits, interventions or other future directions in the research field.

Authors: *We think the reviewer raises an interesting point and agree that cognitive impairment could be associated with reward processing deficits. However, we did not consider analyzing these data owing to the fact that the used paradigm is not appropriate to investigate these associations due to the following reasons: First, due to the pseudo-random block design, each feedback was adapted according to the choice was made by the participants. Thus, every subject achieved exactly the same accuracy rate. Second, since feedback/outcome is independent from reaction times, all subjects were instructed that reaction times were irrelevant for the feedback/amount of money. We now state this more clearly in the paradigm section of our manuscript on page 7. However, we now provide a section for reaction time analyses and a supplemental table for the interested reader (supplemental table 1). The analysis of the reaction times using a 3 (group = BD, UD, HC) × 3 (condition = reward, loss, control) ANOVA with repeated measures yielded a significant main effect of condition ($F_{(2;94)}$=49.4; P<0.01). Post hoc analyses revealed significantly faster reaction times for the control condition compared to the reward ($T_{(96)}$=-8.7; P<0.01) and compared to the loss condition ($T_{(96)}$=-6.8; P<0.01). Furthermore, overall reaction times for the loss condition were slower than for the reward condition ($T_{(96)}$=7.8; P<0.01). Neither*

a main effect of group nor an interaction effect between group and condition was found (P=0.46, see supplement table 1 for details). We further discussed these possible associations within the limitations section on page 16.

2. Similarly, how do the authors view the effect of current mood state and other clinical measures on the BOLD response/connectivity and possibly behaviour?

Authors: *The reviewer points to an interesting question also raised by reviewer 1 and reviewer 2 (see Reviewer #1, Point 2; Reviewer #2, Point 3). Like described above the peak contrast values of the 3 (group) × 2 (condition) interaction analysis of the bilateral NAcc and the potential significant cluster from the functional connectivity analysis were extracted for each participant and further analyzed with PASW Statistics 22 (IBM, Armonk, New York) to investigate the effects of questionnaire measures and clinical variables. An additional analysis of covariance (ANCOVA) was conducted on NAcc responsiveness, with the factor group (UD, BD) as well as BDI, HAMD scores, SHAPS-D scores, medication load index, number of depressive episodes and time since onset of first depressive episode as covariates. This analysis revealed that adding these clinical variables as covariates only slightly weakens the results regarding the condition × group interaction (NAcc right: P = 0.003, NAcc left P = 0.014). Thus, all these covariates seem to have only marginal effects on the reported findings if any. Furthermore, each of these clinical variables was separately correlated with NAcc responsiveness and VTA functional connectivity to explore potential effects. These analyses revealed no significant associations in our sample (all Ps > .11, please see supplemental table 2).*

References

Abler B, Erk S, Walter H (2007). Human reward system activation is modulated by a single dose of olanzapine in healthy subjects in an event-related, double-blind, placebo-controlled fMRI study. *Psychopharmacology* **191**: 823–33.

Franz M, Lemke MR, Meyer T, Ulferts P, Puhl P, Snaith RP (1998). Deutsche Version der Snaith-Hamilton-Pleasure-Scale (SHAPS-D). *Fortschr Neurol Psychiat* **66**: 407–413.

Friston KJ, Buechel C, Fink GR, Morris J, Rolls ET, Dolan RJ (1997). Psychophysiological and modulatory interactions in neuroimaging. *NeuroImage* **6**: 218–229.

Friston KJ, Harrison L, Penny W (2003). Dynamic causal modelling. *NeuroImage* **19**: 1273–302.

Hafeman DM, Chang KD, Garrett AS, Sanders EM, Phillips ML (2012). Effects of medication on neuroimaging findings in bipolar disorder: an updated review. *Bipolar disorders* **14**: 375–410.

Laeger I, Keuper K, Heitmann C, Kugel H, Dobel C, Eden A, *et al* (2014). Have we met before? Neural correlates of emotional learning in women with social phobia. *Journal of Psychiatry & Neuroscience* **39**: E14–23.

Phillips ML, Travis MJ, Kupfer DJ, Fagiolini A (2008). Medication effects in neuroimaging studies of bipolar disorder. *The American Journal of Psychiatry* **165**: 313–320.

Redlich R, Almeida JRC, Grotegerd D, Opel N, Kugel H, Heindel W, *et al* (2014). Brain morphometric biomarkers distinguishing unipolar and bipolar depression: a voxel-based morphometry-pattern classification approach. *JAMA Psychiatry* **71**: 1222–30.

Richard JM, Berridge KC (2013). Prefrontal Cortex Modulates Desire and Dread Generated by Nucleus Accumbens Glutamate Disruption. *Biological Psychiatry* **73**: 360–370.