**Title:**
Discrete neurocognitive subgroups in fully or partially remitted bipolar disorder

**Authors:**
Johan Høy Jensen[a], Ulla Knorr[a], Maj Vinberg[a], Lars V. Kessing[a], Kamilla W. Miskowiak[a, 1]

**Affiliation:**
[a] Psychiatric Centre Copenhagen, Copenhagen University Hospital, Rigshospitalet Dep. 6233, Blegdamsvej 9, DK-2100, Copenhagen, Denmark, Emails: johan.hoey.jensen@regionh.dk, ulla.knorr@regionh.dk, maj.vinberg@regionh.dk, lars.vedel.kessing@regionh.dk, kamilla.miskowiak@regionh.dk

[1] Corresponding author: Kamilla W. Miskowiak, Psychiatric Centre Copenhagen, Copenhagen University Hospital, Rigshospitalet Dep. 6233, Blegdamsvej 9, DK-2100, Copenhagen, Denmark, Tel.: +45 3864 7087, Fax: +45 3864 7077 Email: kamilla@miskowiak.dk

## 1. Introduction

Bipolar disorder (BD) is a severe and chronic illness characterized by recurrent phases of mood swings found in about 3% of the population (Kessler et al., 2005). Although BD has often been regarded as a purely episodic illness, research within the past two decades highlights persistent cognitive and functional impairment in-between the acute mood episodes (Arts et al., 2008; Harvey et al., 2010). Persistent cognitive deficits in remitted patients with BD are well-documented across several cognitive domains including attention, verbal learning, and executive function (Bora et al., 2009; Reichenberg et al., 2009). There is robust evidence from several studies that patients' persistent cognitive dysfunction is a key contributor to their socio-occupational disability independent of mood symptoms (Depp et al., 2012; Martinez-Arán et al., 2007; Mur et al., 2009). Functional impairment is prevalent in BD with unemployment rates substantially higher in BD than the general population as studies show between 4- and 10-fold increase in unemployment among BD patients (Huxley and Baldessarini, 2007; Kogan et al., 2004). Further, approximately two thirds of the patients are unable to regain premorbid levels of social and vocational functioning following a single episode (Huxley and Baldessarini, 2007). A recent meta-analysis found that better verbal memory and executive function were moderately related to positive employment outcome (Tse et al., 2014). This points to cognitive dysfunction as a key treatment priority to improve patients' functional recovery.

Meta-analytic findings indicate that the nature of the persistent cognitive dysfunction in BD is non-specific, involving deficits across several domains with moderate to large effect sizes (Bourne et al., 2013). However, a recent study of N=136 BD patients in full or partial remission suggests that cognitive impairment is not a uniform feature of BD; instead discrete neurocognitive subgroups were found; a "cognitively intact" subgroup with scores equivalent to the performance of healthy control persons (HCs) (31.6%), a subgroup with "selective cognitive impairment" (28.7%), and a subgroup with "global severe impairment" across cognitive domains (39.7%) comparable to

cognitive deficits in schizophrenia (Burdick et al., 2014). However, as the described study (11) is the only published evidence for the presence of neurocognitive subgroups in BD, additional replication studies are needed to make any firm conclusions about the pattern of cognitive dysfunction in BD.

### *1.1 Aims of the study*

The study aims to 1) determine whether there are discrete neurocognitive subgroups in fully or partially remitted patients with BD as prior demonstrated (Burdick et al., 2014). Building onto the findings of Burdick et al. (Burdick et al., 2014), the study aims to 2) examine the pattern of the cognitive deficits and the clinical and functional characteristics of the neurocognitive subgroups in terms of perceived stress, cognitive complaints, work capacity, quality of life, and medication prescribed.

## 2. Materials and methods

### *2.1 Pooling of data*

The present study involved pooling of available baseline data of patients with BD and HCs from three previous studies and an ongoing study. Accordingly, these studies are referred to as study 1 (Miskowia*k et al.* 2014), study 2 (Demant et al., 2015), study 3 (Jensen et al., 2015; Ott et al., 2016), and study 4 (ongoing study) (BD: *N*=201 [study 1=46, study 2=46, study 3=84[2], study 4=25]; HC: *N*=110 [study 3=86, study 4=24]). Eight patients were excluded due to missing data (BD: *N*=193). Data for the HC group came from studies 3 and 4 (*N*=110). We chose to pool data from studies 1-4 because of the largely similar recruitment criteria of BD patients in full or partial remission and large overlap between the applied measures of neurocognitive and functional capacity. Studies 1, 2, and 4 were approved by the Regional Ethics Committee in the Capital Region

[2] 26 of these participants are overlapping with the BD sample in study 4.

of Copenhagen. The local ethics committee stated that there was no need for their approval of study 3.

*2.2 Participants and screening*

A total of 303 individuals aged 18-65 participated in one of the aforementioned studies between September 2009-July 2015, comprising; 193 eligible adult patients with an ICD-10 diagnosis of BD in full or partial remission (defined as Hamilton Depression Rating Scale 17-item [HDRS-17]; ; and the Young Mania Rating Scale [YMRS] scores of ≤7 and 8≤14, respectively) (Hamilton, 1960; Young et al., 1978), and 110 HCs. The HCs were recruited from the blood banks at Copenhagen University Hospital, Rigshospitalet (study 3) or Frederiksberg Hospital (study 4).

The patients were referred by psychiatric specialists in Clinic for Affective Disorders, Psychiatric Centre Copenhagen (studies 1-4) or recruited through website advertisements (study 1).

Common exclusion criteria included prior history of schizophrenia, current substance abuse, substantial somatic illness or a daily use of benzodiazepines ≥22.5 mg oxazepam. Exclusion criteria for the HCs were dyslexia or any personal or family history of mental illness. In studies 1 and 2 additional participation criteria for the BD patients included significant subjective cognitive difficulties according to the Massachusetts General Hospital Cognitive and Physical Functioning Questionnaire (CPFQ; scores ≥4 on ≥2 cognitive domains) (Fav*a et al.* 2009). All participants in studies 2 and 4 were aged 18-50. In study 4, patients had HDRS-17 and YMRS scores of ≤7, respectively, and baseline data were obtained following 2 weeks of remission. Written informed consent was obtained from all participants.

*2.3 Neurocognitive tests*

Overlapping neuropsychological tests from studies 1-4 include: the Trail Making Test part A (TMT-A) (Army Individual Test Battery, 1944), Rey Auditory Verbal Learning Test (RAVLT) (Corwin and Bylsma, 1993; Rey, 1964), Wechsler's Adult Intelligence Scale, 3rd edition (WAIS-III) Letter-Number-Sequencing (Wechsler, 1997), Repeatable Battery for the Assessment of

Neuropsychological Status (RBANS) Digit Span Forward (Randolph, 1998), and verbal fluency (Borkowski et al., 1967) with the letters S and D.

*2.4 Measures of work capacity, subjective cognitive difficulties, perceived stress, and quality of life*

All participants were instructed to complete a set of questionnaires regarding perceived stress (using the 10-item self-report Perceived Stress Scale [PSS]) (Cohen et al., 1983), cognitive and physical functioning (using the 7-item self-report CPFQ) (Fava et al., 2009), work capacities (using the 5-item self-rating Work and Social Adjustment Scale [WSAS]) (Mundt et al., 2002), and quality of life (using the 26-item self-report WHOQOL) (WHOQOL Group, 1994).

*2.5 Statistical analyses*

*2.5.1 Cognitive domain and composite cognition scores*

The neurocognitive test scores of all participants were standardized to *z*-scores (*mean* [*M*] =0, *standard deviation* [*SD*] =1) based on the cognitive performance of the HCs using following formula: (*test score* - *HC test M*) / *HC test SD* (Field, 2013). TMT-A *z*-scores were inversed such that lower scores reflect poorer performance consistent with the direction of the other cognition measures. Extreme *z*-scores (≥4 SDs below the HCs' *M*) were truncated to *z*=-4.0. Using the cognition *z*-scores, we established three cognitive domain *z*-scores for each participant, which again, were standardized (using same formula as above) to the cognitive domain *z*-scores of HCs: processing speed (*z*-scores for TMT-A), verbal learning (averaged *z*-scores for RAVLT: total recall across five learning trials (I-V), recall following interference (IV), and recall following 30 min. delay), and working memory and executive skills (averaged z-scores for WAIS-III Letter-Number Sequencing, RBANS Digit Span Forward, verbal fluency with letters S and D). We chose this grouping of the neuropsychological tests into cognitive domains to be consistent with previous studies (Jensen et al., 2015; Ott et al., 2016). Finally, an overall composite cognition *z*-score was established for each participant by averaging the *z*-scores of the three domains and standardizing this output based on the HCs' composite cognition *z*-scores (same formula as above).

*2.5.2 Establishment and validation of neurocognitive subgroups*

Initially, independent *t*-tests and $\chi^2$-tests were computed between the complete BD and HC samples to detect differences regarding demographical, clinical, cognitive, functional, and medicine characteristics. Then, a hierarchical cluster analysis (HCA) was conducted with the cognitive domain *z*-scores of the complete BD sample to detect homogeneous neurocognitive subgroups. Ward's linkage was selected as agglomeration procedure, and squared Euclidean distance was chosen to compute similarities between cases. The dendogram was visually inspected to detect the optimal number of clusters explaining the cognitive variance. To validate the clusters retained, a discriminant function analysis (DFA) was conducted since this analysis examines the predictive power of each participant's cognitive domain scores to the neurocognitive subgroup. One-way analyses of variance (ANOVA) and least significance difference (LSD) as post-hoc pairwise comparison were applied to adjust for multiple comparison and compare the cognition scores of the BD subgroups and HCs. Cut-off for cognitive impairment was set to scores of $z \leq -1$ (i.e., 1 *SD* or more below the HC *M*). We chose this approach to keep the analyses consistent with the only previously published analyses of neurocognitive subgroups in BD by Burdick et al. (2014).

*2.5.3 Comparisons of neurocognitive subgroups: clinical and functional characteristics*

Post-hoc comparisons of the clinical and functional measures of between the HCA-defined neurocognitive subgroups were conducted to gain insight into the implication of cognitive impairment in BD. In line with previous studies from our group (Demant et al., 2015; Jensen et al., 2015; Ott et al., 2016), the CPFQ cognition measure was created by averaging the *z*-scores for item d (focus), e (recall), and a composite measure composed of *z*-scores for items f (word finding) and g (mental acuity) to ensure that each domain weighted equally in the CPFQ cognition measure. All data for the PSS, CPFQ cognition measure, and WHOQOL were standardized to *z*-scores based on the HCs using the same procedure as for the cognitive data (see above). This was not possible for the WSAS since this questionnaire addresses psychiatric patients only.

Clinical, demographical, and prescribed medicine variables include: gender, age, years of education, onset age, illness duration, depressive episodes and (hypo)-manic episodes (categorization: 1, 2 or ≥3 episodes), medication status (lithium, anticonvulsants, antidepressants, antipsychotics, benzodiazepines, melatonin [yes/no]), total number of medications prescribed, and HDRS-17 and YMRS scores. Between-group comparisons of the neurocognitive subgroups were applied with the aforementioned variables and functional data including the WSAS total scores (i.e., work capacity), PSS total scores (i.e., perceived stress), CPFQ cognition measure (i.e., subjective cognitive difficulties), and WHOQOL (i.e., quality of life) total and domain scores using ANOVA and LSD post-hoc pairwise comparison or $\chi^2$-tests as appropriate.

All statistical analyses were performed using the Statistical Package for Social Sciences (SPSS; version 22.0; IBM Corporation, Armonk, New York). Statistical significance for all tests were set to an alpha-level of $p<.05$ (two-tailed).

## 3. Results

### *3.1 Comparisons of samples*

Comparisons between the complete BD and HC samples showed no differences regarding age, gender or years of education ($p\geq.10$) (for details see Table 1). Comparisons between the BD cohorts from the four original studies revealed subtle differences regarding age, age of onset, illness duration, affective symptoms, TMT-A scores, and medication (LSD: $p$-values$\leq.049$; see Supplementary Table).

### *3.2 Three neurocognitive subgroups*

Visual inspection of the dendogram provided evidence for existence of three neurocognitive subgroups of the complete BD sample (data not shown): one subgroup of 89 patients (46.1%), another subgroup of 63 patients (32.6%) and a third subgroup of 41 (21.2%) patients (see Table 2). The DFA revealed presence of two discriminant functions explaining 77% and 23% of the variance (Wilks' $\lambda=.19$, $\chi^2=317.28$, $p<.01$ and Wilks' $\lambda=.60$, $\chi^2=95.55$, $p<.01$, respectively). The 90% of

original grouped patients were correctly classified with the DFA, indicating the validity of the three neurocognitive subgroups (see Figure 1 for graphical agglomeration of the neurocognitive subgroups).

The ANOVA analyses showed that cognition measures between subgroups and the HCs were significantly different: composite cognition score ($F(3)=99.45$, $p<.01$), verbal learning ($F(3)=36.12$, $p<.01$), working memory and executive skills ($F(3)=28.52$, $p<.01$), and processing speed ($F(3)=140.46$, $p<.01$). Patients in the first (intact) subgroup were cognitively preserved compared to the HCs (range: $z=-.11$ to $.32$). LSD *post-hoc* pairwise comparisons showed no significant differences in verbal learning, working memory and executive skills, and the composite cognition score between the intact patient subgroup and HCs (LSD: $p\geq.19$). Intact patients had significantly higher scores on processing speed than HCs (LSD: $p=.01$). In the second (selectively impaired) subgroup, verbal learning, and working memory and executive skills were within the normal range ($z=-.71$ and $z=-.62$, respectively), but processing speed and, consequently, the overall composite cognition score were severely impaired ($z=-2.53$ and $z=-2.06$, respectively). The selective subgroup scored significantly lower on all cognition measures compared to the intact group (LSD: $p<.01$). The third (globally impaired) group displayed impaired performance on verbal learning ($z=-1.75$), working memory and executive skills ($z=-1.44$), and, consequently, the composite cognition score ($z=-2.02$) while processing speed was within the normal range ($z=-.61$). Post-hoc comparisons between the neurocognitive subgroups showed that all cognition scores of the globally impaired subgroup were significantly lower compared to the intact group (LSD: $p<.01$). Comparison analyses of the global and the selective subgroup revealed greater reduction in processing speed within the selective group (LSD: $p<.01$), and greater impairment of verbal learning and working memory and executive skills in the globally impaired group (LSD: $p$-values$<.01$), while there was no statistical difference in the composite cognition score (LSD: $p\geq.86$) (see Figure 2 and Table 3).

*3.3 Post-hoc validation of the neurocognitive subgroups*

As a *post-hoc* validation of the neurocognitive subgroups identified in the HCA, we conducted multiple regression analyses for the complete BD sample with the three cognitive domains, in turn, as dependent variables. Following variables were entered as independent variables: neurocognitive subgroup (intact/selective/global: 1/2/3), lithium, antipsychotics, and antidepressants (yes/no), benzodiazepines, age, gender, HDRS-17 score, YMRS score, illness duration, number of depressive episodes, manic episodes, number of medications from different classes. The significant predictors for impaired verbal learning ($F(11)=11.52$, $p<.01$, $r^2=.45$, adjusted $r^2=.41$) were categorization in a cognitively impaired subgroup ($\beta=-.55$, $p<.01$) and being male ($\beta=-.32$, $p<.01$) (other p-values $\geq.09$). The significant predictors of working memory and executive dysfunction ($F(11)=9.35$, $p<.01$, $r^2=.39$, adjusted $r^2=.35$) were categorization in a cognitively impaired subgroup ($\beta=-.56$, $p<.01$), more depressive symptoms ($\beta=-.148$, $p=.04$), and a lower number of medications prescribed ($\beta=.21$, $p=.01$) (other $p$-values $\geq.09$). Finally, the only significant predictor for slower processing speed ($F(11)=4.72$, $p<.01$, $r^2=.25$, adjusted $r^2=.20$) was categorization in a cognitively impaired subgroup ($\beta=-.37$, $p<.01$).

*3.4 Comparisons between the BD subgroups: clinical characteristics*

Since we found significant age differences between BD subgroups in the complete patient sample ($F(3)=4.41$, $p<.01$; LSD: $p<.01$), we conducted *post-hoc* comparisons of the BD subgroups' clinical characteristics using analyses of covariance (ANCOVA) with adjustment for age and LSD as pairwise comparison.

No significant differences were found between the three neurocognitive subgroups regarding characteristics, illness load (age of onset, illness duration, mood symptoms, number of previous mood episodes) or BD subtype ($p\geq.064$) (see Table 3). Globally and selectively impaired patients reported poorer work capacity ($F(2)=3.94$, $p=.02$; LSD: $p=.03$ and $p=.01$, respectively), more stress

(LSD: $p<.01$ and $p=.01$, respectively), and more subjective cognitive difficulties (LSD: $p<.01$ and $p=.02$, respectively) than the cognitively intact patients. Selectively impaired patients used a greater number of medications ($F(2)=7.25$, $p=.008$; LSD: $p=.01$ and $p<.01$, respectively), and were more frequently prescribed antidepressants ($F(2)=7.13$, $p=.03$; LSD: $p=.04$ and $p=.02$, respectively) and benzodiazepines ($F(2)=7.13$, $p=.03$; LSD: $p=.04$ and $p<.01$, respectively) relative to the other two neurocognitive subgroups. Moreover, patients in the selectively impaired subgroup had lower quality of life due to the environment than the intact patients ($F(3)=3.38$, $p=.04$; LSD: $p=.02$) (see Table 3).

*3.5 Post-hoc analyses: characteristics of cognitively impaired patients*

For exploratory purposes, we combined the selectively and globally impaired subgroups into one group of 'cognitively impaired' patients who were compared with those who were cognitively intact using ANCOVA adjusted for age. This revealed that cognitively impaired patients had fewer years of education ($F(1)=5.44$, $p=.02$), poorer work capacity (higher WSAS scores: $F(1)=7.92$, $p=.01$), more perceived stress (lower PSS scores: $F(1)=10.00$, $p<.01$), greater subjective cognitive difficulties (higher CPFQ cognition scores: $F(1)=8.30$, $p<.01$), lower quality of life (lower WHOQOL scores: $F(1)=6.71$, $p=.01$), and were prescribed a greater number of medications ($F(1)=7.23$, $p=.01$) compared with the cognitively intact group. These cognitively impaired patients were also more commonly prescribed antipsychotics ($\chi^2(1)=5.43$, $p=.02$) and benzodiazepines ($\chi^2(1)=9.92$, $p<.01$) compared to those who were cognitively intact.

## 4. Discussion

The study investigated the presence and clinical characteristics of discrete neurocognitive subgroups in the, to date, largest cohort of fully or partially remitted BD patients. Three

neurocognitive subgroups were identified; a cognitively intact subgroup (46.1%), a selectively impaired subgroup with severe deficits in processing speed (32.6%), and a globally impaired subgroup with substantial difficulties within verbal learning, working memory, and executive skills (21.2%). Compared to the cognitively intact subgroup, patients with global impairment reported greater stress, more subjective cognitive difficulties, poorer work function, and reduced quality of life in the absence of any differences in other clinical characteristics including age, illness load, or medications. The selectively impaired group also displayed higher levels of stress, more subjective cognitive difficulties, poorer work function, and reduced quality of life compared to the intact group; however, these impairments were generally less pronounced than for globally impaired patients. The selective subgroup was older and was prescribed a greater number of medications, including benzodiazepines and antidepressants relative to the other two neurocognitive subgroups.

Our finding of this large cognitively intact subgroup of BD patients in full or partial remission is in line with previous evidence for an absence of cognitive impairment in 32-42% of patients (Burdick et al., 2014; Reichenberg et al., 2009). Selective severe impairment in processing speed in one third of the BD cohort also corroborates with the previously published findings (Burdick et al., 2014) in a similar albeit smaller cohort of 136 BD patients in full or partial remission (HDRS-17 scores of ≤12, and Clinician-Administered Rating Scale for Mania scores of ≤8) (Altman et al., 1994; Burdick et al., 2014; Hamilton, 1960). Although the present evidence for global impairment in a neurocognitive subgroup of BD patients corresponds to findings of Burdick et al. (Burdick et al., 2014), the proportion of the globally impaired group was substantially smaller (21.2% versus 39.7%) (Burdick et al., 2014). This may be due to differences in the two patient cohorts with our cohort being somewhat younger (*M*[*SD*], 36[10] versus 42[15], respectively) and showing less depressive symptoms (HDRS-17: *M*[*SD*], 6[4] versus 11[9], respectively).

The present and previous (Burdick et al., 2014) evidence for an absence of differences between the globally impaired and intact subgroup in age, severity of affective symptoms, illness load, BD subtype or medication suggests that the observed deficits in verbal memory and executive dysfunction in the globally impaired subgroup cannot be explained by such clinical characteristics. However, we were unable to examine whether cognitive deficits reflect neuroprogression due to lack of detailed information about the exact number of previous affective episodes; 84% and 72% of our BD sample were had experienced at least three previous depressive episodes or (hypo-) manic episodes, respectively. Future prospective studies with long follow-up times are thus needed to clarify the nature and developmental trajectory of neurocognitive deficits in BD.

The selectively impaired subgroup with severe reduction in psychomotor speed was characterized by older age and more polypharmacy than both the intact and globally impaired subgroups. Considerable evidence indicates that mood stabilizing and antipsychotic medication can have negative effects on psychomotor speed (Bora et al., 2009), suggesting that the selectively impaired patients' greater use of medication may to some extent explain their slowed psychomotor speed. However, post-hoc multiple regression analysis of the complete BD sample revealed no significant correlations between slowing of processing speed and polypharmacy, suggesting that the relationship between psychomotor slowing and medication is more complex.

The observation that work capacity, stress, and quality of life were more negatively affected in of the globally and selectively impaired – but otherwise relatively symptom-free – subgroups compared to patients who were cognitively intact, highlights the clinical relevance of targeting cognitive deficits to improve functional recovery of a significant proportion of patients with BD.

Intensive research efforts have recently been made to find new treatments for cognitive dysfunction in BD. Although several pharmacological and psychological interventions are promising (Miskowiak et al., 2014; Torrent et al., 2013), the evidence is still preliminary. A key methodological problem in cognition trials which may partially explain the many negative findings in the field is the general lack of *objective* neuropsychological screening for cognitive deficits before inclusion of participants in these trials. Given our finding that almost half of patients are relatively cognitively intact this is likely to have led to inclusion of a large proportion of patients with little scope for cognitive improvement in these cognition trials. We therefore recommend systematic neuropsychological screening for cognitive dysfunction in BD before commencing a treatment trial targeting cognition to ensure inclusion of an enriched sample of patients with scope for improvement.

A limitation of the study was the cross-sectional study design since this hampers causal inferences regarding the associations between the nature of bipolar disorder itself, cognitive impairment, and functional outcome. Future long-term prospective studies of neurocognitive dysfunction in BD are therefore critically needed. Moderate to severe subjective cognitive difficulties (scores of ≥4 on ≥2 items on the CPFQ) was an inclusion criteria in studies 1 and 2 (i.e., 44% of the present BD sample). This may have introduced an overrepresentation of patients with neurocognitive deficits compared to the general BD population. However, speaking against this, studies by our and other groups have generally found no or only a poor correlation between objective and subjective measures of cognition in remitted BD (Jensen et al., 2015; Rosa et al., 2013). Indeed, we found a larger subgroup of cognitively intact patients than the previously published report (46.1% versus 31.6%) (13). Age differences between the neurocognitive subgroups, may have influenced the HCA detection of three neurocognitive subgroups. Nevertheless, no significant age differences were

found between the globally impaired and the cognitively intact subgroup, and all analyses comparing the clinical characteristics of the neurocognitive subgroups were adjusted for age. The large sample sizes (BD: *N*=193, HC: *N*=110) with same measures of cognition and functional capacity for all participants were strengths of the study, which provide strong statistical power for the presence of the three neurocognitive subgroups and their clinical and functional characteristics.

We identified three neurocognitive subgroups in fully or partially remitted BD patients. Globally and selectively impaired patients displayed more stress, poorer work capacity and reduced quality of life than those who were cognitively intact. These findings highlight a need to screen for cognitive impairment in BD and indicate that novel treatments that target cognition may improve patients' functional recovery.

**Acknowledgements**
The Lundbeck Foundation and TrygFonden are acknowledged for their contributions to KWM's post-doctorate salary at the Psychiatric Center Copenhagen, Copenhagen University Hospital, Rigshospitalet. The Research Fund of the Mental Health Services in the Capital Region of Denmark is acknowledged for UK's post-doctorate salary at the Psychiatric Center Copenhagen, Copenhagen University Hospital, Rigshospitalet.

**References**

1. Altman, E.G., Hedeker, D.R., Janicak, P.G., Peterson, J.L. & Davis, J.M., 1994. The clinician-administered rating scale for mania (CARS-M): development, reliability, and validity. Biol. Psychiat. 36, pp. 124-134.

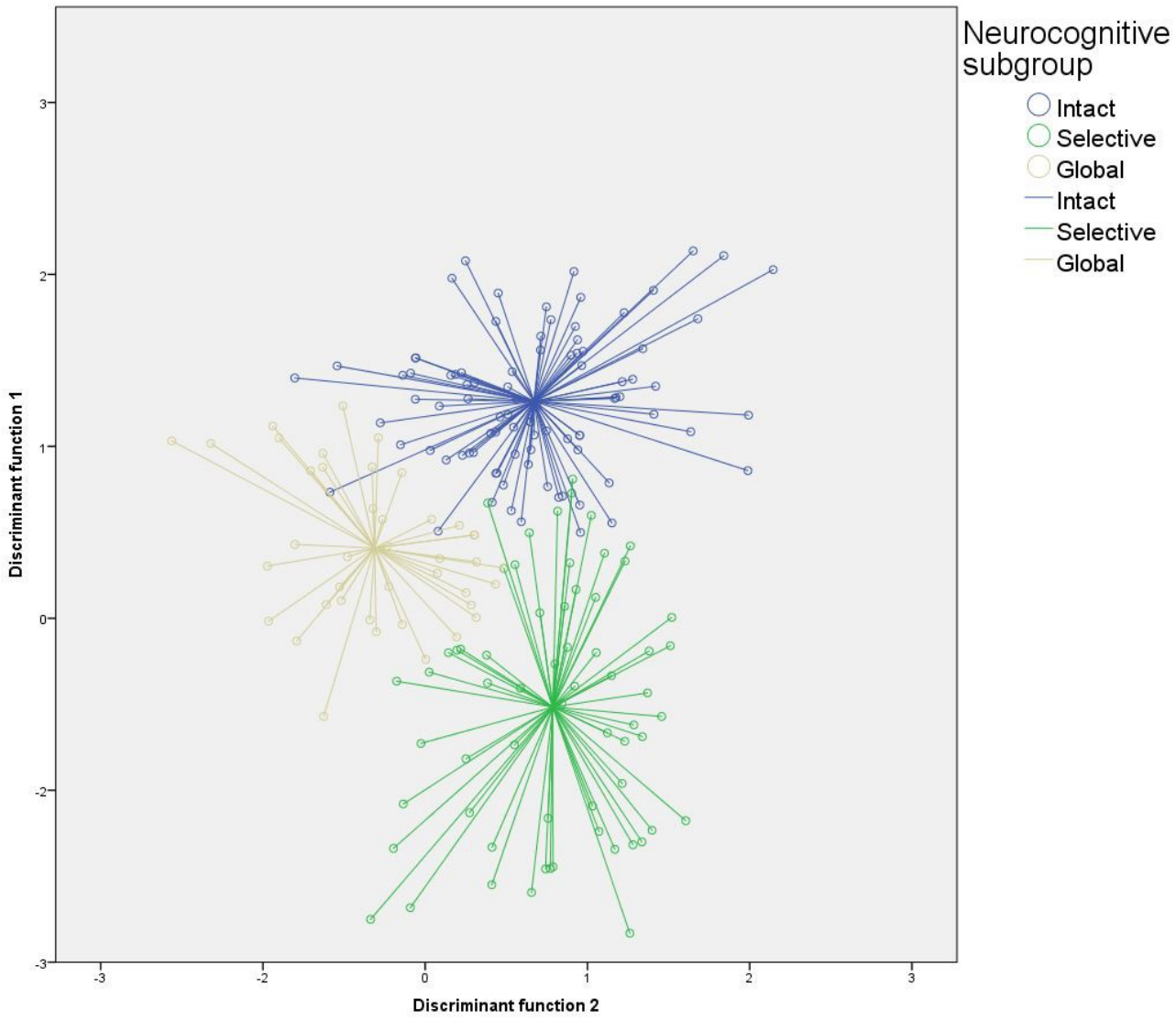
2. Army Individual Test Battery, 1944). Manual of Directions and Scoring. War Department, Adjutant General's Office, Washington, DC.
3. Arts, B., Jabben, N., Krabbendam, L., Van Os, J., 2008. Meta-analyses of cognitive functioning in euthymic bipolar patients and their first-degree relatives. Psychol. Med. 38, pp. 771-785.
4. Bora, E., Yucel, M., Pantelis, C., 2009. Cognitive endophenotypes of bipolar disorder: a meta-analysis of neuropsychological deficits in euthymic patients and their first-degree relatives. J. Affect. Disord. 113, pp. 1-20.
5. Borkowski, J.G., Benton, A.L. & Spreen, O., 1967. Word fluency and brain damage. Neuropsychologia 5, pp. 135-140.
6. Bourne, C., Aydemir, Ö, Balanzá-Martéz, V., Bora, E., Brissos, S., Cavanagh, J.T.O., Clark, L., Cubukcuoglu, Z., Dias, V.V., Dittmann, S., 2013. Neuropsychological testing of cognitive impairment in euthymic bipolar disorder: an individual patient data meta-analysis. Acta Psychiatr. Scand. 128, pp. 149-162.
7. Burdick, K.E., Russo, M., Frangou, S., Mahon, K., Braga, R.J., Shanahan, M., Malhotra, A.K., 2014. Empirical evidence for discrete neurocognitive subgroups in bipolar disorder: clinical implications. Psychol. Med. 44, pp. 3083-3096.
8. Cohen, S., Kamarck, T., Mermelstein, R., 1983. A global measure of perceived stress. J. Health Soc. Behav. 24, pp. 385-396.
9. Corwin, J., Bylsma, F.W., 1993. Psychological examination of traumatic encephalopathy. Clin. Neuropsychol. 7, pp. 3-21.
10. Demant, K., Vinberg, M., Kessing, L.V., Miskowiak, K.M., 2015. Effects of cognitive remediation on cognitive dysfunction in partially or fully remitted individuals with bipolar disorder: results of a randomised controlled trial. PLoS ONE 10, pp. 1-17.
11. Depp, C.A., Mausbach, B.T., Harmell, A. L., Savla, G.N., Bowie, C.R., Harvey, P.D., Patterson, T.L., 2012. Meta-analysis of the association between cognitive abilities and everyday functioning in bipolar disorder. Bipolar Disord. 14, pp. 217-226.
12. Fava, M., Iosifescu, D.V., Pedrelli, P., Baer, L., 2009. Reliability and validity of the Massachusetts general hospital cognitive and physical functioning questionnaire. Psychother. Psychosom. 78, pp. 91-97.
13. Field, A., 2013. Discovering statistics using IBM SPSS statistics. Sage.

14. Hamilton, M., 1960. A rating scale for depression. J. Neurol. Neurosurg. Psychiatr. 23, pp. 56-62.
15. Harvey, P.D., Wingo, A.P., Burdick, K.E. & Baldessarini, R.J., 2010. Cognition and disability in bipolar disorder: lessons from schizophrenia research. Bipolar Disord. 12, pp. 364-375.
16. Huxley, N., Baldessarini, R.J., 2007. Disability and its treatment in bipolar disorder patients. Bipolar Disord. 9, pp. 183-196.
17. Jensen, J.H., Støttrup, M.M., Nayberg, E., Knorr, U., Ullum, H., Purdon, S.E., Kessing, L.V., Miskowiak, K.W., 2015. Optimising screening for cognitive dysfunction in bipolar disorder: Validation and evaluation of objective and subjective tools. J. Affect. Disord. 187, pp. 10-19.
18. Kessler, R.C., Chiu, W.T., Demler, O., Walters, E.E., 2005. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. Arch. Gen. Psychiatr. 62, pp. 617-627.
19. Kogan, J.N., Otto, M.W., Bauer, M.S., Dennehy, E.B., Miklowitz, D.J., Zhang, H., Ketter, T., Rudorfer, M.V., Wisniewski, S.R., Thase, M.E., 2004. Demographic and diagnostic characteristics of the first 1000 patients enrolled in the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD). Bipolar Disord. 6, pp. 460-469.
20. Martinez-Arán, A., Vieta, E., Torrent, C., Sanchez-Moreno, J., Goikolea, J. M., Salamero, M., Malhi, G.S., Gonzalez-Pinto, A., Daban, C., Alvarez-Grandi, S., Fountoulakis, K., Tabares-Seisdesdos, R., Ayuso-Mateos J.L., 2007. Functional outcome in bipolar disorder: the role of clinical and cognitive factors. Bipolar Disord. 9, pp. 103-113.
21. Miskowiak, K.W., Ehrenreich, H., Christensen, E.M., Kessing, L.V., Vinberg, M., 2014. Recombinant human erythropoietin to target cognitive dysfunction in bipolar disorder: a double-blind, randomized, placebo-controlled phase 2 trial. J. Clin. Psychiatr. 75, pp. 1347-1355.
22. Mundt, J.C., Marks, I.M., Shear, M.K., Greist, J. M., 2002. The Work and Social Adjustment Scale: a simple measure of impairment in functioning. Br. J. Psychiatry 180, pp. 461-464.
23. Mur, M., Portella, M.J., Martinez-Arán, A., Pifarre, J., Vieta, E., 2009. Influence of clinical and neuropsychological variables on the psychosocial and occupational outcome of remitted bipolar patients. Psychopathology 42, pp. 148-156.

24. Ott, C.V., Bjertrup, A.J., Jensen, J.H., Ullum, H., Sjælland, R., Purdon, S.E., Vieta, E., Kessing, L.V., Miskowiak, K.W., 2016. Screening for cognitive dysfunction in unipolar depression: Validation and evaluation of objective and subjective tools. J. Affect. Disord. 190, pp. 607-615.
25. Randolph, C., 1998. RBANS manual: Repeatable battery for the assessment of neuropsychological status. The Psychological Corporation, San Antonio, Texas.
26. Reichenberg, A., Harvey, P.D., Bowie, C.R., Mojtabai, R., Rabinowitz, J., Heaton, R.K., Bromet, E., 2009. Neuropsychological function and dysfunction in schizophrenia and psychotic affective disorders. Schizophr. Bull. 35, pp. 1022-1029.
27. Rey, A., 1964. L'examen clinique en psychologie [Clinical psychological examination]. Presses Universitaires de France, Paris.
28. Rosa, A.R., Mercade, C., Sanchez-Moreno, J., Sole, B., Bonnin, C.D.M., Torrent, C., Grande, I., Sugranyes, G., Popovic, D., Salamero, M., Kapczinski, F., Vieta, E., Martinez-Arán, A., 2013. Validity and reliability of a rating scale on subjective cognitive deficits in bipolar disorder (COBRA). J. Affect. Disord. 150, pp. 29-36.
29. Torrent C., Bonnin C.D.M., Martínez-Arán A., Valle J, Amann B.L., González-Pinto A., Crespo J.M., Ibáñez Á., Garcia-Portilla M.P., Tabarés-Seisdedos R., Arango C., Colom F., Solé B., Pacchiarotti I., Rosa A.R., Ayuso-Mateos J.L., Anaya C., Fernández P., Landín-Romero R., Alonso-Lana S., Ortiz-Gil J., Segura B., Barbeito S., Vega P., Fernández M., Ugarte A., Subirà M., Cerrillo E., Custal N., Menchón J.M., Saiz-Ruiz J., Rodao J.M., Isella S., Alegría A., Al-Halabi S., Bobes J., Galván G., Saiz P.A., Balanzá-Martínez V., Selva G., Fuentes-Durá I., Correa P., Mayoral M., Chiclana G., Merchan-Naranjo J., Rapado-Castro M., Salamero M., Vieta E., 2013. Efficacy of functional remediation in bipolar disorder: a multicenter randomized controlled study. American Journal of Psychiatry .
30. Tse, S., Chan, S., Ng, K.L., Yatham, L.N., 2014. Meta-analysis of predictors of favorable employment outcomes among individuals with bipolar disorder. Bipolar Disord. 16, pp. 217-229.

31. Wechsler, D., 1997. WAIS-III: Administration and scoring manual: Wechsler adult intelligence scale. Psychological Corporation, San Antonio, Texas.
32. WHOQOL Group, 1994. The development of the World Health Organization quality of life assessment instrument (the WHOQOL). Springer, Berlin, Heidelberg, pp. 41-57.
33. Young, R.C., Biggs, J.T., Ziegler, V.E., Meyer, D.A., 1978. A rating scale for mania: reliability, validity and sensitivity. Br. J. Psychiatry 133, pp. 429-435.

**Figure 1**
Graphical agglomeration of the neurocognitive subgroups.

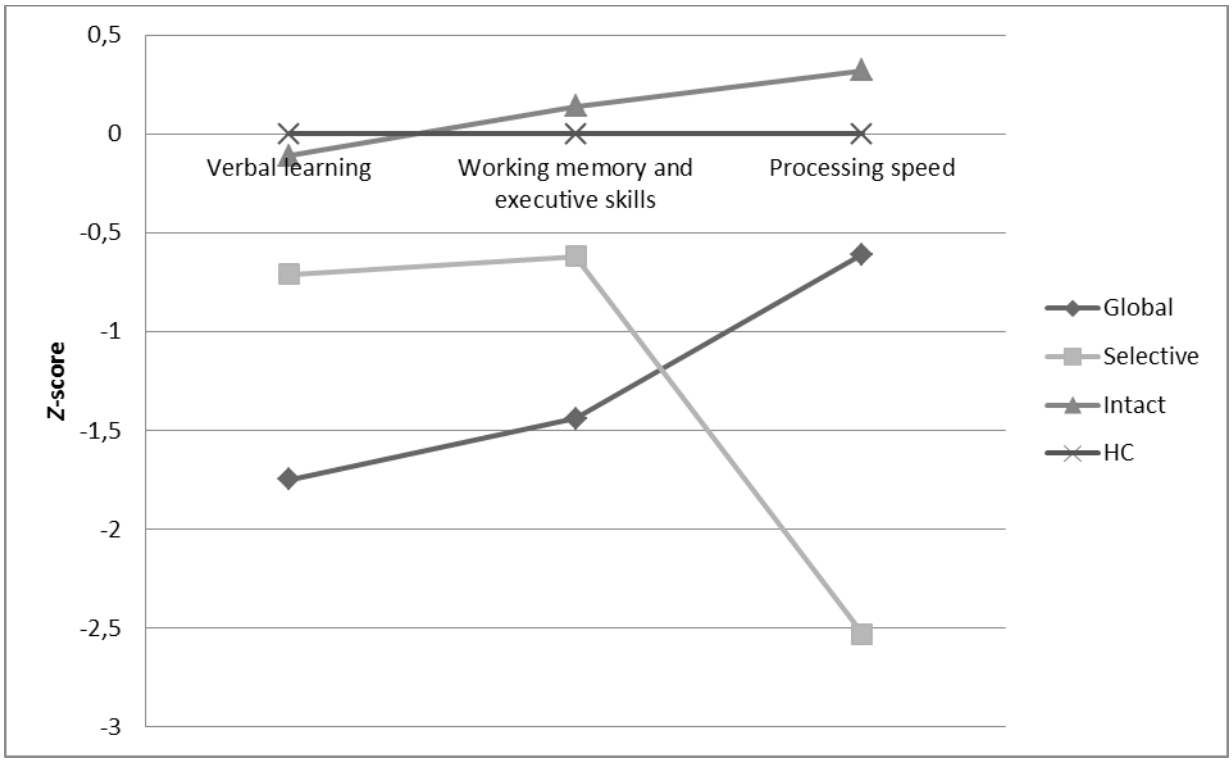



Scatter-plot and centroids of the discriminant values from the discriminant function analysis based

on the neurocognitive subgroups established from the hierarchical cluster analysis.
Abbreviations: Global = globally impaired subgroup, Selective = selectively impaired subgroup,

Intact = cognitively intact subgroup, HC = healthy control

**Figure 2**
Neurocognitive profiles of the three subgroups and HC sample.




The X-axis indicates the cognitive domains and the Y-axis indicate the mean cognition *z*-score of the neurocognitive subgroups and the HC sample.
Abbreviations: Global = globally impaired subgroup, Selective = selectively impaired subgroup, Intact = cognitively intact subgroup, HC = healthy control

**Table 1**
Characteristics of the complete BD and HC samples.

| | | | Statistics | | |
|---|---|---|---|---|---|
| | BD patients | HC | *df* | *t* or $\chi^2$ | *p* |
| *N* | 193 | 110 | | | - |
| Age | 36 (10) | 35 (12) | 200.05 | .97 | .335 |
| Gender, F/M (%) | 62/38 | 57/43 | 1 | .70 | .401 |
| Years of education | 15 (3) | 16 (2) | 300 | -1.65 | .101 |
| Age of onset | 22 (10) | - | | | - |

| | | | | | |
|---|---|---|---|---|---|
| Illness duration, years | 15 (11) | - | | | - |
| HDRS-17 | 6 (4) | 1 (1) | 300 | 11.07 | <.001 |
| *Full remission, ≤7, no. (%)* | 132 (69) | 110 (100) | | | - |
| *Partial remission, 8≤14 no. (%)* | 60 (31) | 0 (0) | | | - |
| YMRS | 3 (3) | 1 (1) | 299 | 6.21 | <.001 |
| *Full remission, ≤7, no. (%)* | 177 (93) | 110 (100) | | | - |
| *Partial remission, 8≤14 no. (%)* | 14 (7) | 0 (0) | | | - |
| BD type, I/II (%) | 58/42 | - | | | - |
| Depressive episodes (1/2/>3), (%) | 2/12/84 | - | | | - |
| (Hypo-) manic episodes (1/2/>3), (%) | 11/17/72 | - | | | - |
| Composite cognition score | -1.0 (1.5) | 0.0 (1.0) | 271.03 | -6.68 | <.001 |
| *Verbal learning and memory* | -.66 (1.2) | 0.0 (1.0) | 255.75 | -5.17 | <.001 |
| *Working memory and executive skills* | -.44 (1.2) | 0.0 (1.0) | 300 | -3.32 | .001 |
| *Processing speed* | -.80 (1.5) | 0.0 (1.0) | 294.42 | -5.56 | <.001 |
| WSAS | 16.5 (8.6) | - | | | - |
| PSS | -2.3 (1.8) | 0.0 (1.0) | 298.71 | -14.44 | <.001 |
| CPFQ cognition measure | -4.5 (3.2) | 0.0 (1.0) | 220.60 | -18.91 | <.001 |
| WHOQOL | -3.1 (1.3) | 0.0 (1.0) | 266.50 | -23.50 | <.001 |
| Current medication | | | | | |
| *Number of medications* | 2.1 (1.0 | - | | | - |
| *Lithium, no. (%)* | 110 (58) | - | | | - |
| *Anticonvulsants, no. (%)* | 118 (62) | - | | | - |
| *Antidepressants, no. (%)* | 40 (21) | - | | | - |
| *Antipsychotics, no. (%)* | 89 (47) | - | | | - |
| *Benzodiazepines, no. (%)* | 33 (17) | - | | | - |
| *Melatonin, no. (%)* | 5 (7) | - | | | - |
| *No medication, no. (%)* | 3 (2) | - | | | - |

Independent samples t-test and $\chi^2$ as appropriate. Measures of cognition, the PSS, CPFQ cognition measure, and WHOQOL are given as z-scores, mean (standard deviation), based on the HCs' performance.

Abbreviations: BD = Bipolar disorder, HC = Healthy control, HDRS-17 = Hamilton Depression Rating Scale 17-item, YMRS = Young Mania Rating Scale, WSAS = Work and Social Adjustment Scale, PSS = Perceived Stress Scale, CPFQ = Massachusetts General Hospital Cognitive and Physical Functioning Questionnaire, WHOQOL = World Health Organization's Quality of Life.

**Table 2**
Comparison between the BD subgroups and HCs across cognitive domain and composite cognition scores.

| | Global (*n*=41, 21.2%) | Selective (*n*=63, 32.6%) | Intact (*n*=89, 46.1%) | HC (*n*=110) | *df* | *F* | *p* |
|---|---|---|---|---|---|---|---|
| Composite cognition score | –2.02 (.66) | –2.06 (1.14) | .20 (.92) | 0.0 (1.0) | 3 | 99.45 | <.001<br>HC *v*. G, *p*<.001<br>HC *v*. S, *p*<.001<br>HC *v*. I, *p*=.187<br>G *v*. S, *p*=.860<br>G *v*. I, *p*<.001<br>S *v*. I, *p*<.001 |
| Verbal learning | –1.75 (.77) | –.71 (1.11) | -.11 (.97) | 0.0 (1.0) | 3 | 36.12 | <.001<br>HC *v*. G *p*<.001<br>HC *v*. S *p*<.001<br>HC *v*. I *p*=.438<br>G *v*. S *p*<.001<br>G *v*. I *p*<.001<br>S *v*. I *p*<.001 |
| Working memory and executive skills | –1.44 (.62) | –.62 (1.03) | .14 (1.12) | 0.0 (1.0) | 3 | 28.52 | <.001<br>HC *v*. G *p*<.001<br>HC *v*. S *p*<.001<br>HC *v*. I *p*=.336<br>G *v*. S *p*<.001<br>G *v*. I *p*<.001<br>S *v*. I *p*<.001 |
| Processing speed | –.61 (.78) | –2.53 (1.12) | .32 (.63) | 0.0 (1.0) | 3 | 140.46 | <.001<br>HC *v*. G *p*<.001 |

| | |
|---|---|
| | HC *v*. S *p*<.001 |
| | HC *v*. I *p*=.013 |
| | G *v*. S p<.001 |
| | G *v*. I *p*<.001 |
| | S *v*. I *p*<.001 |

Analyses of variance with least significance differences as pairwise comparison. Measures of cognition are given as z-scores, mean (standard deviation), based on the HCs' performance.
Abbreviations: BD = bipolar disorder, HC = healthy control, df = degrees of freedom, G = globally impaired subgroup, S = selectively impaired subgroup, I = cognitively intact subgroup.

**Table 3**
*Comparison of the characteristics of the neurocognitive cognitive subgroups.*

| | Subgroup | | | Statistics | | |
|---|---|---|---|---|---|---|
| | Global | Selective | Intact | *df* | *F* or $\chi^2$ | *p* |
| *n* | 41 | 63 | 89 | | | |
| Age, *min-max* | 38 (11), 20-65 | 39 (11), 18-58 | 34 (9), 19-61 | 2 | 6.70 | .002 |
| | | | | | | G *v.* I, *p*=.020 |
| | | | | | | S *v.* I, *p*=.001 |
| Gender, F/M (%) | 66/34 | 64/36 | 60/40 | 2 | .54 | .762 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Years of education | 14 (3) | 15 (3) | 15 (3) | 2 | 2.80 | .064 |
| Age of onset | 23 (10) | 21 (11) | 21 (9) | 2 | 1.10 | .334 |
| Illness duration, years | 15 (10) | 18 (11) | 13 (9) | 2 | 1.10 | .334 |
| HDRS-17 | 6 (5) | 6 (4) | 5 (4) | 2 | 1.74 | .178 |
| YMRS | 3 (3) | 3 (3) | 2 (3) | 2 | .16 | .856 |
| BD type, I/II (%) | 61/39 | 61/39 | 53/47 | 2 | 1.17 | .557 |
| Depressive episodes (1/2/≥3), (%) | 0/13/87 | 2/12/86 | 8/11/81 | 4 | 5.62 | .230 |
| (Hypo-) manic episodes (1/2/≥3), (%) | 8/14/78 | 11/20/69 | 13/16/71 | 4 | 1.39 | .847 |
| WSAS | 18 (9) | 18 (9) | 14 (8) | 2 | 3.94 | .021 |
| | | | | | | G *v.* I, *p*=.030 |
| | | | | | | S *v.* I, *p*=.014 |
| PPS | –2.84 (1.89) | –2.62 (1.59) | –1.89 (1.81) | 2 | 5.16 | .007 |
| | | | | | | G *v.* I, *p*=.005 |
| | | | | | | S *v.* I, *p*=.013 |
| CPFQ cognition measure | –8.04 (4.95) | –6.86 (4.00) | –5.49 (4.53) | 2 | 5.00 | .008 |
| | | | | | | G *v.* I, *p*=.003 |
| | | | | | | S *v.* I, *p*=.049 |
| WHOQOL | –3.40 (1.32) | –3.35 (1.14) | –2.89 (1.29) | 2 | 3.38 | .038 |
| | | | | | | G *v.* I, *p*=.042 |
| | | | | | | S *v.* I, *p*=.027 |
| *Physical health* | –3.56 (1.55) | –3.38 (1.48) | –2.98 (1.62) | 2 | 2.73 | .068 |
| *Psychological health* | –2.48 (1.35) | –2.76 (1.48) | –2.35 (1.62) | 3 | 2.10 | .126 |
| *Social relationsships* | –1.32 (1.66) | –1.45 (1.31) | –1.03 (1.42) | 3 | .81 | .447 |
| *Environment* | –2.38 (1.21) | –2.48 (1.36) | –2.01 (1.26) | 3 | 3.38 | .036 |
| | | | | | | S *v.* I, *p*=.015 |
| Current medication | | | | | | |
| *Number of medications* | 2.1 (0.9) | 2.6 (1.0) | 1.9 (1.0) | 2 | 7.25 | .001 |
| | | | | | | G *v.* S, *p*=.009 |
| | | | | | | S *v.* I, *p*<.001 |
| *Lithium, no. (%)* | 25 (61) | 37 (60) | 48 (55) | 2 | .64 | .727 |

| | | | | | | |
|---|---|---|---|---|---|---|
| *Anticonvulsants, no. (%)* | 23 (56) | 41 (66) | 54 (61) | 2 | 1.01 | .587 |
| *Antidepressants, no. (%)* | 6 (15) | 20 (32) | 14 (16) | 2 | 7.13 | .028 |
| | | | | 1 | 4.06 | G *v.* S, *p*=.044 |
| | | | | 1 | 5.55 | S *v.* I, *p*=.019 |
| *Antipsychotics, no. (%)* | 23 (56) | 33 (53) | 33 (38) | 2 | 5.51 | .064 |
| *Benzodiazepines, no. (%)* | 6 (15) | 20 (32) | 7 (8) | 2 | 15.29 | <.001 |
| | | | | 1 | 4.06 | G *v.* S, *p*=.044 |
| | | | | 1 | 14.56 | S *v.* I, *p*<.001 |
| *Melatonin, no. (%)* | 0 (0) | 0 (0) | 5 (6) | 2 | 6.01 | .050 |
| *No medication, no. (%)* | 1 (2) | 0 (0) | 2 (2) | 2 | 1.47 | .480 |

Analyses of covariance adjusted for age and least significance difference as pairwise comparison and $\chi^2$ as appropriate. Measures of cognition, the PSS, CPFQ cognition measure, and WHOQOL are given as z-scores, mean (standard deviation), based on the HCs' performance.

Abbreviations: HC = healthy controls, df = degrees of freedom, HDRS-17 = Hamilton Depression Rating Scale 17-item, YMRS = Young Mania Rating Scale, BD = bipolar disorder, WSAS = Work and Social Adjustment Scale, PSS = Perceived Stress Scale, CPFQ = Massachusetts General Hospital Cognitive and Physical Functioning Questionnaire, WHOQOL = World Health Organization's Quality of Life, G = globally impaired subgroup, S = selectively impaired subgroup, I = cognitively intact subgroup.

**Abstract**

*Background:* Neurocognitive impairment in remitted patients with bipolar disorder contributes to functional disabilities. However, the pattern and impact of these deficits are unclear.

*Methods:* We pooled data from 193 fully or partially remitted patients with bipolar disorder and 110 healthy controls. Hierarchical cluster analysis was conducted to determine whether there are discrete neurocognitive subgroups in bipolar disorder. The pattern of the cognitive deficits and the characteristics of patients in these neurocognitive subgroups were examined with analyses of covariance and least significance difference pairwise comparison.

*Results:* Three discrete neurocognitive subgroups were detected: one that was cognitively intact (46.1%), one that was selectively impaired with deficits only in processing speed (32.6%), and one that was globally impaired across verbal learning, working memory, and executive skills (21.2%). The globally and selectively impaired subgroups were characterized by greater perceived stress and cognitive complaints, poorer work capacity and reduced quality of life compared to patients who were cognitively intact.

*Limitations*: The study design was cross-sectional which limits inferences regarding the causality of the findings.

*Conclusion:* Globally and selectively impaired bipolar disorder patients displayed more functional disabilities than those who were cognitively intact. The present findings highlight a clinical need to systematically screen for cognitive dysfunction in remitted bipolar disorder and to target residual cognitive dysfunction in future treatment strategies.

**Highlights**

Three subgroups with unique neurocognitive profiles were found in bipolar disorder.

Intact cognitive function was found in 46% patients.

Subgroups with global (21%) or selective (33%) deficits had most functional problems.

There is need for a more tailored treatment targeting cognition in bipolar disorder.

**Supplementary Table.**

Comparison of the characteristics of the BD samples from the four original studies 1-4.

| | **Pooled BD samples** | | | | **Statistics** | | |
|---|---|---|---|---|---|---|---|
| | Study 1 | Study 2 | Study 3 | Study 4 | *df* | *F* or $\chi^2$ | *p* |
| *n* | 46 | 38 | 84 | 25 | | | |
| Age | 34 (8) | 41 (12) | 36 (10) | 33 (10) | 3 | 4.48 | .005 |
| | | | | | | | S1 *v.* S2, *p*=.001 |
| | | | | | | | S2 *v.* S3, *p*=.014 |
| | | | | | | | S2 *v.* S4, *p*=.003 |
| Gender, F/M (%) | 65/35 | 66/34 | 64/36 | 44/56 | 3 | 4.06 | .255 |
| Years of education | 16 (3) | 15 (3) | 15 (3) | 15 (2) | 3 | 1.05 | .370 |
| Age of onset | 17 (7) | 23 (10) | 25 (11) | 16 (6) | 3 | 11.76 | <.001 |
| | | | | | | | S1 *v.* S2, *p*=.002 |
| | | | | | | | S1 *v.* S3, *p*<.001 |
| | | | | | | | S2 *v.* S4, *p*=.002 |
| | | | | | | | S3 *v.* S4, *p*<.001 |
| Illness duration, years | 17 (8) | 18 (11) | 11 (10) | 18 (11) | 3 | 5.68 | .001 |
| | | | | | | | S1 *v.* S3, *p*=.002 |
| | | | | | | | S2 *v.* S3, *p*=.001 |
| | | | | | | | S3 *v.* S4, *p*=.008 |
| HDRS-17 | 6 (4) | 9 (4) | 5 (4) | 3 (2) | 3 | 12.39 | <.001 |
| | | | | | | | S1 *v.* S2, *p*=.001 |
| | | | | | | | S1 *v.* S4, *p*=.003 |
| | | | | | | | S2 *v.* S3, *p*<.001 |
| | | | | | | | S2 *v.* S4, *p*<.001 |
| YMRS | 2 (3) | 3 (3) | 3 (4) | 1 (1) | 3 | 2.91 | .036 |
| | | | | | | | S2 *v.* S4, *p*=.041 |
| | | | | | | | S3 *v.* S4, *p*=.024 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| BD type, I/II (%) | 65/34 | 42/58 | 61/39 | 56/44 | 3 | 5.24 | .155 |
| Depressive episodes (1/2/≥3), (%) | 2/11/87 | 0/9/91 | 6/16/78 | 8/4/88 | 6 | 6.74 | .345 |
| (Hypo-) manic episodes (1/2/≥3), (%) | 9/12/79 | 11/17/72 | 16/16/68 | 0/28/72 | 6 | 7.65 | .265 |
| Cognitive composite score | -1.08 (1.23) | -1.51 (1.46) | -.79 (1.59) | -.84 (1.29) | 3 | 2.30 | .079 |
| *Verbal learning and memory* | -.78 (1.01) | -.81 (1.05) | -.50 (1.21) | -.74 (1.42) | 3 | .96 | .413 |
| *Working memory and executive skills* | -.58 (1.02) | -.57 (1.06) | -.30 (1.33) | -.50 (1.03) | 3 | .81 | .489 |
| *Processing speed* | -.68 (1.30) | -1.46 (1.63) | -.71 (1.58) | -.35 (1.22) | 3 | 3.50 | .017 |
| | | | | | | | S1 *v.* S2, *p*=.017 |
| | | | | | | | S2 *v.* S3, *p*=.010 |
| | | | | | | | S2 *v.* S4, *p*=.004 |
| Current medication | | | | | | | |
| *Number of medications* | 2.4 (.9) | 2.1 (.8) | 2.2 (1.2) | 1.7 (.9) | 3 | 2.89 | .037 |
| | | | | | | | S1 *v.* S4, *p*=.005 |
| | | | | | | | S3 *v.* S4, *p*=.029 |
| *Lithium, no. (%)* | 32 (70) | 14 (37) | 52 (63) | 12 (28) | 3 | 11.48 | .009 |
| | | | | | 1 | 9.00 | S1 *v.* S2, *p*=.003 |
| | | | | | 1 | 7.41 | S2 *v.* S3, *p*=.006 |
| *Anticonvulsants, no. (%)* | 35 (76) | 20 (53) | 48 (59) | 15 (60) | 3 | 5.73 | .125 |
| *Antidepressants, no. (%)* | 13 (28) | 19 (50) | 8 (10) | 0 (0) | 3 | 33.69 | <.001 |
| | | | | | 1 | 4.17 | S1 *v.* S2, *p*=.041 |
| | | | | | 1 | 7.36 | S1 *v.* S3, *p*=.007 |
| | | | | | 1 | 8.65 | S1 *v.* S4, *p*=.003 |
| | | | | | 1 | 24.12 | S2 *v.* S3, p<.001 |
| | | | | | 1 | 17.90 | S2 *v.* S4, *p*<.001 |
| *Antipsychotics, no. (%)* | 24 (52) | 8 (21) | 46 (56) | 11 (44) | 3 | 13.58 | .004 |
| | | | | | 1 | 8.55 | S1 *v.* S2, *p*=.003 |
| | | | | | 1 | 12.89 | S2 *v.* S3, *p*<.001 |
| | | | | | 1 | 3.77 | S2 *v.* S4, *p*=.049 |
| *Benzodiazepines, no. (%)* | 6 (13) | 14 (37) | 11 (13) | 2 (8) | 3 | 13.12 | .004 |
| | | | | | 1 | 6.50 | S1 *v.* S2, *p*=.011 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | 1 | 8.64 | S2 *v.* S3, *p*=.003 |
| | | | | | 1 | 6.62 | S2 *v.* S4, *p*=.010 |
| *Melatonin, no. (%)* | 1 (2) | 2 (5) | 1 (1) | 1 (4) | 3 | 1.90 | .594 |
| *No medication, no. (%)* | 0 (0) | 0 (0) | 2 (2) | 1 (4) | 3 | 2.70 | .441 |

*ANOVA with least significance differences as pairwise comparison and $\chi^2$ as appropriate. Measures of cognition are given as z-score mean (standard deviation) relative to the HCs' performance.*

Abbreviations: S1 = study 1 (Miskowiak et al., 2014), S2 = study 2 (Demant et al., 2015), S3 = Study 3 (Jensen et al., 2015; Ott et al., 2016), S4 = study 4 (ongoing study), df = degress of freedom, HDRS-17 = Hamilton Depression Rating Scale 17-item, YMRS = Young Mania Rating Scale, BD = bipolar disorder