
Copyright:
This is the authors’ accepted manuscript of an article published in its final definitive form by Cambridge University Press, 2015.

Link to published article:
http://dx.doi.org/10.1017/S0033291715000926

Date deposited:
08/10/2015

Embargo release date:
15 December 2015

This work is licensed under a [Creative Commons Attribution-NonCommercial 3.0 Unported License](http://creativecommons.org/licenses/by-nc/3.0/).
Neurocognitive intra-individual variability in mood disorders: effects on attentional response time distributions

Peter Gallagher *, 1, BSc(Hons), MPhil, PhD, AFBPsS
Jonna Nilsson 1,5, BSc(Hons), PhD
Andreas E. Finkelmeyer 1, MSci, PhD
Michelle Goshawk 1, BSc(Hons), MSc
Karine A. Macritchie 2, MD, MBChB, MRCPsych
Adrian J. Lloyd 1,4, MBBS, MD (Hons), MRCPsych
Jill M. Thompson 1, BSc (Hons), PhD, CPsychol, CSci, AFBPsS
Richard J. Porter 3, MBBS, MD, MA, DipObst, MRCPsych
Allan H. Young 2, MBChB, MPhil, PhD, FRCPsych, FRCP(C)
I. Nicol Ferrier 1, BSc(Hons), FRCP, MD(Hons), FRCPsych
R. Hamish McAllister-Williams 1,4, BSc(Hons), MBChB, MD, PhD, FRCPsych
Stuart Watson 1,4, MD, MBBS, MRCPsych

1 Institute of Neuroscience, Newcastle University, UK
2 King’s College London, Institute of Psychiatry, UK
3 Department of Psychological Medicine, University of Otago, Christchurch, New Zealand
4 Northumberland, Tyne and Wear NHS Foundation Trust, UK
5 Ageing Research Institute, Karolinska Institute, Sweden

* Address for Correspondence:
Dr Peter Gallagher, Institute of Neuroscience, Newcastle University, The Henry Wellcome Building,
Framlington Place, Newcastle upon Tyne, NE2 4HH, UK

Tel.: +44 (0)191 208 7166 Fax.: +44 (0)191 208 5227 Email: peter.gallagher@newcastle.ac.uk

Abstract: 244 words. Manuscript: 4,100 words, with 2 tables and 2 figures.

This work was supported by grant funding from the Stanley Medical Research Institute (REF: 03T-429) and the Medical Research Council (REF: GU0401207). PG, INF, RHMcW and SW received Research Capability Funding from the Northumberland, Tyne and Wear NHS Foundation Trust.
Abstract

**Background.** Attentional impairment is a core cognitive feature of major depressive disorder (MDD) and bipolar disorder (BD). However, little is known of the characteristics of response time (RT) distributions from attentional tasks. This is crucial to furthering our understanding of the profile and extent of cognitive intra-individual variability (IIV) in mood disorders.

**Method.** A computerised sustained attention task was administered to 138 healthy controls and 158 patients with a mood disorder: 86 euthymic BD, 33 depressed BD and 39 medication-free MDD patients. Measures of IIV, including individual standard deviation (iSD) and coefficient of variation (CoV), were derived for each participant. Ex-Gaussian (and Vincentile) analyses were used to characterise the RT distributions into three components: mu and sigma (mean and standard deviation of the Gaussian portion of the distribution) and tau (the ‘slow tail’ of the distribution).

**Results.** Compared to healthy controls, iSD was increased significantly in all patient samples. Due to minimal changes in average RT, CoV was only increased significantly in BD depressed patients. Ex-Gaussian modelling indicated a significant increase in tau in euthymic BD (Cohen’s d=0.39,95%CI=0.09-0.69; p=0.011), and both sigma (d=0.57,95%CI=0.07-1.05; p=0.025) and tau (d=1.14,95%CI=0.60-1.64; p<0.00001) in depressed BD. The mu parameter did not differ from controls.

**Conclusions.** Increased cognitive variability may be a core feature of mood disorders. This is the first demonstration of differences in attentional RT distribution parameters between MDD and BD, and BD depression and euthymia. These data highlight the utility of applying measures of IIV to characterise neurocognitive variability and the great potential for future application.

**Key words:** neuropsychology; attention; variability; ex-Gaussian; bipolar disorder; major depression.
Introduction

Neurocognitive dysfunction is a common feature of mood disorders. Deficits in a range of cognitive processes have been described during symptomatic episodes in major depressive disorder (MDD) (Zakzanis et al. 1998; Lee et al. 2012; Rock et al. 2014) and bipolar disorder (BD) (Rubinsztein et al. 2006; Kurtz & Gerraty 2009; Gallagher et al. 2014; Gallagher et al. 2015), including in medication-free patients (Porter et al. 2003; Taylor Tavares et al. 2007). There has long been an emphasis on the extent to which such deficits can be observed in clinical remission (Astrup et al. 1959; Bratfos & Haug 1968), with growing consensus that they may be state-independent (Robinson et al. 2006; Torres et al. 2007; Arts et al. 2008; Bora et al. 2012; Bourne et al. 2013). The further identification – albeit less consistently – of modest dysfunction in the non-affected, first-degree relatives of affected probands (Balanzá-Martínez et al. 2008; Bora et al. 2009) has resulted in some aspects of neurocognitive dysfunction being put forward as candidate cognitive endophenotypes for mood disorders. Due to a paucity of studies in some areas, there remains debate over the extent to which specific cognitive deficits can be viewed as true endophenotypes (i.e. heritable, co-segregating, and found in non-affected family members at a higher rate than in the general population; Gottesman & Gould 2003) rather than core illness ‘traits’, emerging consequent to the mood disorder (Glahn et al. 2004; Christensen et al. 2006; Daban et al. 2012).

Impairments in facets of attentional processing have been described in many studies of neurocognitive function in mood disorders (Cohen et al. 2001). Deficits have been observed in MDD and BD patients when euthymic (Paelecke-Habermann et al. 2005; Torrent et al. 2006; Preiss et al. 2009; Robinson et al. 2013) as well as abnormalities in the activation of underlying neurocircuitry when performing attentional tasks (Strakowski et al. 2004; Mullin et al. 2012). Following the observation of deficits in first-degree relatives of BD patients, and euthymic recurrent MDD patients, attentional control (cognitive flexibility) has been suggested as candidate endophenotype for mood disorder in general (but not actual disease phenotypes) (Clark et al. 2005b). One of the most frequently examined aspects of attention in mood disorders has been vigilance (or sustained attention). Performance decrements, which increase with time-on-task, on the degraded stimulus form of the continuous performance test (CPT) in euthymic BD
patients have led to the suggestion that alterations in sustained attention may be an endophenotype for BD (Ancín et al. 2010). Numerous other studies have demonstrated CPT deficits in BD and MDD in symptomatic states (Koetsier et al. 2002; Porter et al. 2003; Fleck et al. 2012; Gallagher et al. 2014) and in euthymia (Wilder-Willis et al. 2001; Liu et al. 2002; Weiland-Fiedler et al. 2004; Doyle et al. 2005; Thompson et al. 2005; Kolur et al. 2006). CPT deficits have also been observed in some (Klimes-Dougan et al. 2006; Trivedi et al. 2008) but not all (Clark et al. 2005a; Meyer & Blechert 2005; Jabben et al. 2009; Walshe et al. 2012) studies in first-degree relatives. A recent study found both behavioural deficits and functional magnetic resonance imaging (fMRI) differences (increased activation in the insula and parts of the cingulate cortex) during a CPT in euthymic BD-I patients and non-affected relatives compared to controls (Sepede et al. 2012).

One important consideration in the assessment of attentional processes is in the method of performance measurement. In most CPTs, absolute errors, signal detection indices or mean reaction time (RT) over sub-components or the overall task are typically used. However, increasingly there is recognition of the need to go beyond such measures and take into account inconsistency of responses or *intra-individual variability* (IIV). This can be achieved most simply by calculation of the standard deviation of item-by-item RT for each individual (or the individual standard deviation; *iSD*), although as this measure is strongly related to mean RT, the *coefficient of variation* (CoV) is often preferred (Jackson et al. 2012) which divides the *iSD* by the corresponding individual’s mean RT. Such measures are being increasingly applied in the cognitive ageing literature (Nilsson et al. 2014), where it has been reported that IIV indices are better than mean RT in differentiating early neurodegeneration from healthy aging (Hultsch et al. 2002), and are strongly related to broader cognitive function (Bielak et al. 2010) and brain white matter integrity (Fjell et al. 2011; Jackson et al. 2012). However, empirical RT distributions are fundamentally non-normal and tend to be positively skewed and there is growing interest in the utility of mathematical RT modelling to characterise dissociable components of RT distributions (Balota & Yap 2011).
The *ex-Gaussian distribution*, a mathematical convolution of a Gaussian (normal) and exponential distribution, produces a good approximation to empirical RT distributions (Schmiedek *et al.* 2007). The ex-Gaussian distribution has three parameters: *mu* and *sigma*, the mean and standard deviation of the Gaussian (normal) component; and *tau*, which determines the exponential component and represents the relative strength of the ‘slow-tail’ of the distribution (Ratcliff 1979). As the ex-Gaussian model represents the distribution of RT, it can intuitively be related to ‘standard’ arithmetic properties, for example, the sum of *mu* and *tau* equals the overall arithmetic mean of the data (Ratcliff 1979; Heathcote *et al.* 1991). This methodology has been used to model RT in a number of attentional tasks in older adults, for example, demonstrating a clear increase in the *tau* component in mild dementia of the Alzheimer’s-type compared to controls, which correlated with decreased cerebral white matter (Tse *et al.* 2010; Jackson *et al.* 2012). More generally, RT variability has been linked to white matter integrity across the normal developmental trajectory in healthy children, adolescents and adults: maturation of white matter integrity and connectivity leading to reductions in RT IIV (Fjell *et al.* 2011; Tamnes *et al.* 2012). Given the growing evidence of impaired white matter integrity in MDD and BD and those at high-risk (Heng *et al.* 2010; Macritchie *et al.* 2010; Sprooten *et al.* 2011; Henderson *et al.* 2013; Leow *et al.* 2013; Sarrazin *et al.* 2014; Wang *et al.* 2014) there is a clear rationale for applying such analyses to attentional RT data in mood disorder.

Despite the potential utility of these approaches, there is very little data on IIV in mood disorders. Increased variability on the Connors CPT in manic and euthymic patients has been reported (Bora *et al.* 2006), although variability was examined between average blocks of trials rather than individual RT. One study found a large effect size in the increase in RT *iSD* from a CPT in young BD probands and their unaffected first-degree relatives compared to matched controls (Brotman *et al.* 2009). It has been reported that RT *iSD* from a Go/No-go paradigm was increased in patients with schizophrenia/schizoaffective disorder, but not in those with major depression or borderline personality disorder compared to healthy controls (Kaiser *et al.* 2008). To date there has been no comprehensive assessment of attentional IIV, with full RT modelling, in mood disorders.
The aim of the present study was therefore to examine RT distributions from an attentional CPT in patients with mood disorders, comparing $iSD$, CoV and ex-Gaussian components ($\mu$, $\sigma$ and $\tau$) in patients with bipolar disorder (euthymia and depression), medication-free depression and healthy control participants. As the ex-Gaussian is a parametric model of an underlying theoretical distribution, Vincentile analysis was also conducted in order to demonstrate convergence across the two techniques (Tse et al. 2010). This non-parametric technique directly assesses raw empirical RT distributions and makes no assumptions about an underlying theoretical distribution (by first ordering and then dividing the empirical distribution into a number of equal-sized ‘bins’ and computing the average RT in each of these bins). It was hypothesised that, overall, the mood disorder groups would show a significantly increased IIV and ex-Gaussian $\tau$ component (reflecting increased response variability, especially slowing) compared to matched controls.

**Methods and Materials**

Individual RT datasets were collated from multiple studies conducted in the Institute of Neuroscience (Academic Psychiatry), Newcastle University which had used the same attentional task (Porter et al. 2003; Thompson et al. 2005; Macritchie et al. 2010; Gallagher et al. 2014).

**Participants**

Patients aged 18 to 65 years with a diagnosis of bipolar disorder (BD) confirmed using the Structured Clinical Interview for DSM-IV (SCID; First et al. 1995), were recruited from secondary and tertiary care services in North East of England. All were out-patients and either currently in a depressive episode (SCID defined) or euthymic, prospectively defined as $\leq 7$ on both the Hamilton Depression Rating Scale (HDRS21; Hamilton 1960) and the Young Mania Rating Scale (YMRS; Young et al. 1978) at initial assessment and after 1 month. Patients were excluded if they met criteria for any other current Axis I disorder (except anxiety) or substance dependence/abuse. All were receiving medication at the time of testing but this had remained stable for $\geq 4$ weeks. For the MDD cohort, patients aged 18 to 65 years with
a DSM–IV diagnosis of major depressive disorder (MDD), single episode or recurrent, were recruited from general practice clinics. For this latter (MDD) cohort, patients had been entirely psychotropic medication-free for at least 6 weeks before recruitment and were excluded if currently taking other medication active in the central nervous system, including beta-blockers or St. John’s Wort, or if there was a comorbid medical/psychiatric diagnosis, or recent alcohol/substance misuse. All were tested as soon as possible after recruitment to minimise delay in treatment. For all participants, illness characteristics, clinical ratings and medication history were determined by trained psychiatrists using full history, case-note and medication review and standardized rating scales. All studies were approved by the local NHS Research Ethics Committee and all participants gave written, informed consent.

Neurocognitive testing

All participants completed the Vigil continuous performance test (Cegalis & Bowlin 1991) using the same parameters. In this task, a continuous stream of random letters of the English alphabet is displayed on a computer screen. Each letter appears for 85ms, followed by a 900ms inter-stimulus interval (ISI) and is presented as a white letter on a black background in the centre of the screen (see Figure 1). Participants are instructed to look out for a target sequence (an ‘A’ immediately followed by a ‘K’) and must respond “as quickly, but as accurately as possible” by pressing the spacebar if this target sequence occurs. The letter ‘A’ thereby becomes the signal for the potential occurrence of a target sequence, but responses should only be made once the second letter of the sequence, ‘K’, appears. In total, 480 letters are displayed, in which 100 target sequences occur. These are pseudo-randomised between each quarter of the test i.e. so there are 25 targets within every 120 trials (The Psychological Corporation 1998).

Data analysis procedure

Data extraction and cleaning

RT data were re-extracted from the original Vigil CPT output files and any responses were either classified as ‘valid’ or as ‘commission error’ according to their temporal relationship to the target.
Response times were always measured in relation to the onset of the second stimulus of a target sequence (letter ‘K’). In contrast to the standard analysis, we classified responses as ‘valid’ even if they occurred after the onset of the letter that immediately followed a target sequence (see Figure 1), allowing maximum response times of up to 1970ms (i.e. \( [2 \times ISI] + [2 \times \text{letter duration}] \)). However, there is one exception to this rule: since it is possible that two (or more) target sequences follow directly after another (i.e. ‘A-K-A-K’), responses to the second ‘A’ would no longer be considered valid for the initial target sequence, as such a response could be a premature response to the new target sequence. Such responses were classified as commission errors. Any other responses that could not be associated with a target stimulus according to the above rules were also classified as commission errors. Target stimuli with no detectable valid response were classified as ‘misses’.

Insert figure 1 about here

This classification scheme ensured that responses with RT just above the ISI were considered (late) valid responses to the target, instead of resulting, according to the original scheme, in a “miss” classification to the target stimulus and a commission error for the stimulus following the target. While we believe that this classification better reflects the underlying psychological processes, it is important to consider the number of misses when looking at the distribution of response times of an individual. For instance, some individuals may have been better able than others to withhold responses when they detected that those responses would be late (i.e. after the onset of the stimulus following a target), thereby restricting their maximum response times to the “standard” response window. Since such behaviour would reduce the potential range of RTs and therefore RT variability, care must be taken that this reduction does not come at a cost of an increased number of misses.

---

1 this was done to permit the analysis of RT in relation to the intended target, independent of ISI. In typical analysis of continuous attention tasks, the RT is limited to a maximum ≤ ISI ms. For example, if a participant is slow to recognise given target sequences and make a response, even though their responses may be initiated validly by targets, they will be incorrectly recorded as errors if a subsequent letter is presented before their response can be completed. Most often these will appear as very fast commission errors.
**IIV analysis and ex-Gaussian modelling**

From valid responses, basic measures of IIV were derived using the $iSD$ – the SD of all RT for each individual, and the $CoV$ – the $iSD$ divided by an individual’s mean RT. Ex-Gaussian probability density functions were fitted to the distribution of valid response times of each individual using the DISTRIB toolbox (Lacouture & Cousineau 2008) in MATLAB® v.R2010b (The MathWorks Inc. 2010). This toolbox uses maximum likelihood principles to estimate the ex-Gaussian distribution parameters $mu$, $sigma$ and $tau$. Vincentile plots were also derived as a distribution-free representation of the data. For these data, RTs within each participant were ranked and eight Vincentiles derived (representing the average RT within each sequential 12.5% of valid data, from fastest to slowest). Individual Vincentiles were then averaged across participants.

**Healthy control reference data**

An SAS algorithm was used (Kosanke & Bergstralh 1995) which sampled from the overall control cohort (n=138) and matched controls to individual cases according to age, sex and NART estimated IQ (Nelson 1982). This created very closely matched healthy control groups for each of the three patient groups. Group analyses were made using SPSS v19.

**Results**

**Subject demographics and clinical details**

In total 297 datasets were available for analysis (see table 1). This included 138 healthy controls (61 males, 77 females) and 159 patients. The three patient samples included: 86 euthymic bipolar patients (41 males, 45 females), 33 depressed bipolar patients (19 males, 14 females) and 39 depressed MDD patients (15 males, 24 females). Data from one further female depressed MDD patient were excluded from the analysis as only 22% valid responses were recorded for this patient. The three patient groups and their respective matched control groups were closely matched for age and NART-estimated IQ ($p>0.69$ for all).
None of the patients in the MDD group were currently on psychotropic medication. Twenty four (62%) had never previously taken antidepressant medication; of the remaining 15 (38%), the median time since last treatment was 12 months (range 2-84 months). Five bipolar patients were drug-free at the time of testing. In the euthymic sample, n=76 (88%) were taking a mood stabiliser (of which n=55 lithium), n=23 (27%) antidepressant medication, and n=23 (27%) antipsychotic medication. In the depressed sample, n=27 (84%) were taking a mood stabiliser (of which n=8 lithium), n=26 (81%) antidepressant medication, and n=15 (47%) an antipsychotic medication. Medication details of one patient were not recorded.

**Response profiles**

Within the original raw dataset (n=296), a total of 29,677 individual trials were recorded, of which 28,482 (96.0%) were responses within the originally defined response window (0-985ms). The remaining 4.0% were classified as: early (300/29,677; 1.0%), i.e. responses that occurred before the “K” of an “AK” target sequence; or late (201/29,677; 0.7%) i.e. ‘correct’ responses which were slow (985-1970ms); or misses (694/29,677; 2.3%). Examining these between patients and controls indicated that the greater proportion of early (226/300; 75.3%) and late responses (152/201; 75.6%), and misses (570/694; 82.1%) occurred in the patient sample. Comparing these directly revealed that, on average, significantly more misses occurred in all 3 patient samples compared to their respective control group, with depressed BD patients also making more early and late responses (see table 2).

Following data cleaning (see methods above), an average of 94.4 (SD=8.14) responses per participant in patients and 98.3 (SD=3.11) responses per participant in controls were available for RT analysis.

---

2 As this method of classification recoded the majority of what would previously have been considered ‘commission errors’ into ‘correct-late’ responses, in the present analysis commission errors were very infrequent and not considered further.
Average RT

The analysis of the standard average RT showed significantly slower RT for the group of euthymic BD patients ($F_{1,170}$=6.322, $p=0.013$; $d=0.38$, 95%CI=0.08 to 0.68; see Table 2), but not for the group of depressed BD patients ($F_{1,64}$=1.009, $p=0.319$; $d=0.25$, 95%CI=-0.24 to 0.73) or the group of depressed MDD patients ($F_{1,76}$=0.048, $p=0.826$; $d=0.05$, 95%CI=-0.39 to 0.49) compared to controls.

IIV indices

The various measures of intraindividual RT variability are shown in Table 2. Analysis of the iSD demonstrated significantly greater variability in patients compared to their matched control data, for euthymic BD ($F_{1,170}$=4.785, $p=0.030$; $d=0.33$, 95%CI=0.03 to 0.63), depressed BD ($F_{1,64}$=32.474, $p<0.00001$; $d=1.40$, 95%CI=0.85 to 1.92) and depressed MDD ($F_{1,76}$=5.662, $p=0.020$; $d=0.54$, 95%CI=0.08 to 0.99).

Accounting for the overall mean RT, a significantly greater CoV was observed in depressed BD ($F_{1,64}$=28.824, $p<0.00001$; $d=1.32$, 95%CI=0.77 to 1.84). There was also a statistical trend for greater CoV for depressed MDD ($F_{1,76}$=3.545, $p=0.064$; $d=0.43$, 95%CI=-0.02 to 0.87), but no difference in euthymic BD ($F_{1,170}$=0.732, $p=0.393$; $d=0.13$, 95%CI=-0.17 to 0.43).

Ex-Gaussian analysis and Vincentile Plots

The ex-Gaussian analysis indicated that there were differences across the 3 distribution parameters (see table 2). No significant differences between patients and controls were observed in mu (euthymic BD: $F_{1,170}$=1.943, $p=0.165$; $d=0.21$, 95%CI=-0.09 to 0.51; depressed BD: $F_{1,64}$=1.864, $p=0.177$; $d=-0.34$, 95%CI=-0.82 to 0.15; depressed MDD: $F_{1,76}$=0.301, $p=0.585$; $d=-0.12$, 95%CI=-0.57 to 0.32). No significant differences in the sigma parameter were observed for euthymic BD ($F_{1,170}$=1.918, $p=0.168$; $d=0.21$, 95%CI=-0.09 to 0.51) or depressed MDD ($F_{1,76}$=1.901, $p=0.172$; $d=0.31$, 95%CI=-0.14 to 0.76), but sigma was significantly increased in depressed BD ($F_{1,64}$=5.292, $p=0.025$; $d=0.57$, 95%CI=0.07 to 1.05). A significant increase in the exponential part of the RT distribution was observed for both BD patient groups: the tau parameter was greater in euthymic BD ($F_{1,170}$=6.604, $p=0.011$; $d=0.39$, 95%CI=-0.09 to
and depressed BD ($F_{1,64}=21.347$, $p<0.0001$; $d=1.14$, 95%CI=0.60 to 1.64) compared to controls. There was also a statistical trend for greater $\tau$ in depressed MDD ($F_{1,76}=3.034$, $p=0.086$; $d=0.39$, 95%CI=-0.06 to 0.84).

Vincentile plots are shown in figure 2, providing convergent support for the ex-Gaussian analyses. For the euthymic BD sample, the plots for patients are controls remain close until the last Vincentile ($V_8$) where they diverge more sharply. This occurs more clearly in the depressed MDD and BD samples, particularly the latter. However, there are also differences evident in the first Vincentile ($V_1$) for the depressed samples, with responses being faster in patients than controls (a difference which is significant in the BD depressed sample ($p=0.024$).

To facilitate comparison between patient groups, the ex-Gaussian parameters for euthymic BD, depressed BD and MDD groups were expressed as a $z$-score based on the mean and SD of their respective control groups. One-way ANOVA revealed significant differences for $\mu$ ($F_{2,155}=4.348$, $p=0.015$) and $\tau$ ($F_{2,155}=15.545$, $p<0.0001$). Post hoc contrasts revealed that the $\mu$ parameter was significantly different between euthymic and depressed BD groups ($p=0.006$) with a trend between euthymic BD and MDD groups ($p=0.085$). For $\tau$, the depressed BD group differed significantly from both euthymic and MDD groups ($p<0.001$) (see table 2 for data).

**Receiver Operating Characteristic (ROC) analysis**

To demonstrate the degree of differentiation between the clinical groups and controls (i.e. that differences are not consequent to extreme responses from a small number of participants), an ROC plot (Wilcoxon estimate) was used to determine the optimum cut-point to maximise sensitivity and specificity. For MDD, a $\tau$ value of 56.12 yielded a ROC AUC=0.60 (95%CI=0.46 to 0.73), with sensitivity=0.74 and specificity=0.44. For euthymic BD, a $\tau$ value of 56.35 yielded a ROC AUC=0.62 (95%CI=0.53 to 0.70), with sensitivity=0.77 and specificity=0.44. For depressed BD, a $\tau$ value of 85.56
yielded a ROC AUC=0.82 (95%CI=0.70 to 0.93), with sensitivity=0.73 and specificity=0.88. Comparing between the clinical groups, the \( \tau \) parameter also differentiated depressed from euthymic BD patients with sensitivity=0.70 and specificity=0.71 (ROC AUC=0.73, 95%CI=0.61 to 0.84), and depressed BD from depressed MDD with sensitivity=0.73 and specificity=0.65 (ROC AUC=0.68, 95%CI=0.55 to 0.82).

**Relationship to severity of depression**

Exploratory Spearman’s correlations were performed separately for each patient group, between IIIV parameters and the HDRS21. No significant correlations between \( iSD, \) CoV or ex-Gaussian parameters were observed in euthymic (\(-0.073 \leq r_s \leq 0.188, \ p > 0.080\) for all) or depressed BD (\(-0.135 \leq r_s \leq -0.017, \ p > 0.450\) for all). For MDD patients, a near-significant positive correlation between depression severity and CoV was observed (\(r=0.314, \ p = 0.051\)).

**Discussion**

The present study investigated intra-individual RT variability during a simple sustained attention task in three groups of patients with mood disorders, euthymic BD, depressed BD and depressed MDD. All three groups showed evidence of increased response variability compared to matched controls. Euthymic BD patients had greater values of \( iSD \) and \( \tau \), but not in CoV or \( \sigma \). Together with the fact that this group also showed greater standard average RT, but not in the fitted \( \mu \) parameter, these results indicate that the differences between these patients and controls is best characterized as in increase in the exponential part of the RT distribution (i.e. an increased number of ‘disproportionately slow’ responses), as this would cause a shift in mean RT and \( iSD \) but not in CoV. Depressed BD patients showed the most consistent evidence of increase in RT variability, as all four indices of variability (\( iSD, \) CoV, \( \sigma \) and \( \tau \)) were significantly increased in comparison to the healthy control sample. It may at first seem surprising that there was no significant increase in average RT in this group as a result of increased variability. However, as can be seen in the Vincentile plot of this group, the increase in variability was due not only to an increase in the number of slow responses (similar to euthymia), but also the number...
of fast responses (although not to a sufficient extent to alter \( \mu \)). Depressed MDD patients showed the weakest evidence for a RT variability increase. While the \( iSD \) was significantly higher in this group, both the \( CoV \) and the \( \tau \) parameter showed only statistical trends for larger values. There were no differences in average RT, \( \mu \) or \( \sigma \).

These data are in line with previous reports of increased IIV in attentional performance in BD (Bora et al. 2006; Brotman et al. 2009). However, to our knowledge this is the first paper to comprehensively examine RT distribution parameters and IIV across patients with mood disorders. Previous studies have applied ex-Gaussian RT modelling to tasks in children and adolescents with ADHD. The \( \tau \) parameter has been suggested to produce excellent differentiation between ADHD and controls (Leth-Steensen et al. 2000). Subsequent findings suggest that there are differences in all three parameters compared to controls, with more variability (\( \sigma \)) and increases in \( \mu \) and particularly slow (\( \tau \)) responses – the latter suggested to reflect attentional lapses in some but not all trials (Hervey et al. 2006). In the present study, while there was no significant difference in \( \mu \) between patients and controls, the Vincentile plots did indicate some evidence of faster responses in \( V_1 \) in the depressed samples (which was significant in BD depression). There was also a significant increase in the number of misses in all patient groups (and early and late responses, in depressed BD), compared to controls. This general inconsistency combined with the frequency of disproportionately slow responses is again consistent with ‘phasic’ attentional task engagement/disengagement. This has been suggested previously during CPT task performance in euthymic BD (Robinson et al. 2013). Functional imaging has further revealed that while prefrontal activation occurs early during CPT performance in mania, it cannot be maintained over sustained periods (Fleck et al. 2012).

An area of ongoing debate is the extent to which RT distribution characteristics can be linked to specific aspects of neurocognitive function. For example, the utility of ex-Gaussian modelling has been demonstrated across different conditions of the classic Stroop test, revealing attentional shifts which would otherwise be missed with outcomes based on simple central tendency (Heathcote et al. 1991).
These authors suggest that no direct attribution can be made between ‘parameter and process’ and while “the ex-Gaussian model describes RT data successfully, it does so without the benefit of an underlying theory” (Heathcote et al. 1991). However, more recently it has been proposed that the tau parameter is strongly related to ‘higher’ cognitive functions (a statistical composite measure of working memory tasks and reasoning) and is therefore a marker of individual differences in attentional/executive control (Schmiedek et al. 2007). As work in this area progresses – and if IIV and ex-Gaussian measures are applied more frequently in clinical studies – it may be possible to derive more precise theoretical accounts, informing our understanding of neurocognition in mood disorders.

A strength of the present study was the assessment of IIV and application of RT modelling to one single attentional CPT which had been used consistently in a series of studies in the same research centre. However, it should be noted that in addition to attention, other cognitive processes such as processing speed have been assessed as putative cognitive endophenotypes in BD (Antila et al. 2011; Daban et al. 2012). One caveat is that most studies have used the Digit-Symbol task as an index of processing speed, but this measure is known to involve multiple interacting lower-level and higher-level cognitive control processes, including executive control and attention (Cepeda et al. 2013). Therefore when utilising such tasks in the search for candidate endophenotypes, especially if proposing process-specificity, it is necessary to consider more precisely the cognitive processes underpinning performance on any given measure. It is also important to ascertain whether IIV and shifts in the RT distribution in mood disorders are sensitive to the demand characteristics of tasks, such as rate of presentation or cognitive load, and therefore whether they are related more to impairments in attentional control or basic processing efficiency.

Other methodological considerations should be highlighted. The present study utilised a large normative reference sample from which control data was selected by computer algorithm and demographically-matched to individual patient cases. This ensured very close group-wise matching of patients and controls which was independent of experimenter selection. The majority of BD patients in the present
study were taking psychotropic medication at the time of testing. While several studies have reported minimal effects of medication on performance (Goswami et al. 2009; Bourne et al. 2013), the potential impact of medication on performance should be considered and replication in medication-free samples or in cohorts large enough to perform sub-group analysis is needed. The depressed MDD sample in the present study was entirely psychotropic medication-free at the time of testing and some evidence of increased IIV was observed, specifically $iSD$, but the ex-Gaussian parameters were not significantly different from controls (although $\tau$ was increased at a trend level, with a small-medium effect size). Differences in clinical characteristics (see Porter et al. 2003), such as medication, age (the MDD patients were younger) and number of episodes (the majority of MDD patients being first-episode) mean that comparisons need to be interpreted cautiously. Similarly the inherent difficulty in how to equate stage of illness and other clinical characteristics between MDD and BD in order to reliably compare them should also be noted, along with the issue of statistical power in relation to the sample size characteristics.

The clearest comparison between IIV parameters can be made between the BD groups. It is of note that variability is evident in euthymia (as increased $iSD$ and $\tau$) but increases in depression, reflected in the additional increase in $CoV$ and $sigma$. It would be of interest for future studies to explore the potential neurobiological mechanisms underlying such effects. For example, it has been demonstrated in animal and human models that corticosteroid (cortisol) levels can exert both positive and negative effects on attention, depending on the relative occupancy of corticosteroid receptors (Lupien & McEwen 1997). Given the evidence of hypothalamic-pituitary-adrenal (HPA) axis dysfunction and hypercortisolaemia in BD (Rybakowski & Twardowska 1999; Gallagher et al. 2007), which is present in euthymia but worse in depression, examining the hypothesized role of systems such as the HPA axis and their potential for causing or exacerbating state-related effects is warranted.

Due to the methodological issues outlined it remains to be established if specific features of cognitive processes, such as intra-individual variability in sustained attention, could be considered as cognitive endophenotypes. It has previously been suggested that impairment on tasks such as the CPT is more an
indicator of general brain dysfunction, underpinning the attentional system, than a disorder-specific marker (Rosvold et al. 1956; Riccio et al. 2002). Given the strong relationship that has been identified between IIV and white matter, it is possible that some measures of IIV or components of the RT distribution such as tau, are sensitive markers of general white matter integrity (Fjell et al. 2011; Jackson et al. 2012; Tamnes et al. 2012). These links warrant detailed exploration in future studies – especially in combination with focussed processing speed and attentional assessment – to ascertain the utility of these measures as markers of structural and functional integrity in a variety of clinical disorders in which white matter impairments are implicated, such as neurodegenerative and mood disorders (Sachdev et al. 2005; Assareh et al. 2011; Poletti et al. 2015). Including assessment in individuals with genetic risk, for example for mood disorder, will further inform the extent to which they can be considered endophenotypic markers (Hasler et al. 2006). Developing understanding of the relationship between specific cognitive processes and their structural and functional underpinnings has clear clinical implications, especially in the potential use of neurocognitive function in the stratification of mood disorders (Insel et al. 2010).

The present study has demonstrated increased RT IIIV in sustained attention in mood disorders. Further analysis of RT distribution parameters revealed differences in the parameters affected between MDD and BD, and depression and euthymia in BD. These data highlight the utility of applying measures of IIIV to characterise cognitive variability and the potential for future application in studies examining neurocognitive dysfunction and its underlying functional and structural brain connectivity in mood disorder.
Acknowledgments

We are grateful to the participants who contributed to the research and to those clinicians involved in the wider research programme, including recruitment and screening: Niraj Ahuja, Sankalpa Basu, Jane Carlile, Louise Golightly, Thiyyancheri Harikumar, Patrick Keown, Samer Makhoul, Anuradha Menon, Gavin Mercer, Rajesh Nair, Bruce Owen, and Nanda Palanichamy.

This work was supported by grant funding from the Stanley Medical Research Institute (REF: 03T-429) and the Medical Research Council (REF: GU0401207). PG, INF, RHMcW and SW received Research Capability Funding from the Northumberland, Tyne and Wear NHS Foundation Trust.

Declaration of Interest

None.
References


Table 1. Demographic and clinical details

<table>
<thead>
<tr>
<th></th>
<th>All healthy controls (n=138)</th>
<th>BD euthymic (n=86)</th>
<th>Control comparison a (n=86)</th>
<th>BD depressed (n=33)</th>
<th>Control comparison a (n=33)</th>
<th>MDD (n=39)</th>
<th>Control comparison a (n=39)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>Age</td>
<td>40.5</td>
<td>12.54</td>
<td>44.0</td>
<td>9.74</td>
<td>43.8</td>
<td>9.61</td>
<td>47.0</td>
</tr>
<tr>
<td>NART estimated IQ</td>
<td>111.1</td>
<td>8.64</td>
<td>110.9</td>
<td>10.28</td>
<td>111.4</td>
<td>8.63</td>
<td>109.0</td>
</tr>
<tr>
<td>HAMD21</td>
<td>-</td>
<td>-</td>
<td>1.6</td>
<td>1.66</td>
<td>-</td>
<td>-</td>
<td>21.9</td>
</tr>
<tr>
<td>Age of onset</td>
<td>-</td>
<td>-</td>
<td>24.8</td>
<td>7.12</td>
<td>-</td>
<td>-</td>
<td>25.8</td>
</tr>
<tr>
<td>Bipolar I or II (n)b</td>
<td>-</td>
<td>-</td>
<td>70 BD-I, 16 BD-II</td>
<td>-</td>
<td>-</td>
<td>12 BD-I, 16 BD-II</td>
<td>-</td>
</tr>
</tbody>
</table>

a Each control comparison was sampled from the overall control group (see methods) so are not independent.

b SCID diagnosed bipolar type I or II (missing for n=5 BD depressed)
Table 2. Descriptive statistics for RT data and response profile

<table>
<thead>
<tr>
<th></th>
<th>All healthy controls (n=138)</th>
<th>BD euthymic (n=86)</th>
<th>Control comparison a (n=86)</th>
<th>BD depressed (n=33)</th>
<th>Control comparison a (n=33)</th>
<th>MDD (n=39)</th>
<th>Control comparison a (n=39)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean  SD</td>
<td>Mean  SD</td>
<td>Mean  SD</td>
<td>Mean  SD</td>
<td>Mean  SD</td>
<td>Mean  SD</td>
<td>Mean  SD</td>
</tr>
<tr>
<td>Average RT (ms)</td>
<td>375.9  69.08</td>
<td>411.0*  75.56</td>
<td>382.7  71.91</td>
<td>412.9  96.07</td>
<td>390.9  80.84</td>
<td>382.7  88.91</td>
<td>378.3  87.33</td>
</tr>
<tr>
<td>iSD</td>
<td>83.6  29.78</td>
<td>95.9*  29.93</td>
<td>85.4  33.34</td>
<td>143.8***  56.21</td>
<td>80.7  29.70</td>
<td>104.8*  50.28</td>
<td>83.1  26.53</td>
</tr>
<tr>
<td>CoV</td>
<td>0.23  0.08</td>
<td>0.24  0.07</td>
<td>0.23  0.09</td>
<td>0.36***  0.14</td>
<td>0.21  0.08</td>
<td>0.27  0.09</td>
<td>0.23  0.09</td>
</tr>
<tr>
<td>Ex-Gaussian parameters</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mu</td>
<td>310.0  76.60</td>
<td>332.2  76.55</td>
<td>316.0  75.93</td>
<td>296.0  87.66</td>
<td>324.6  82.76</td>
<td>298.5  78.66</td>
<td>309.4  95.24</td>
</tr>
<tr>
<td>Sigma b</td>
<td>32.1  20.82</td>
<td>37.7  21.19</td>
<td>33.2  21.78</td>
<td>45.2*  33.85</td>
<td>29.8  18.46</td>
<td>40.8  39.62</td>
<td>31.2  17.97</td>
</tr>
<tr>
<td>Tau</td>
<td>66.0  29.14</td>
<td>78.8*  32.55</td>
<td>66.8  28.98</td>
<td>117.3***  59.40</td>
<td>66.3  22.32</td>
<td>84.5  47.67</td>
<td>68.9  28.71</td>
</tr>
<tr>
<td>Response profile c</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early response</td>
<td>0.54  1.09</td>
<td>0.70  1.22</td>
<td>0.62  1.29</td>
<td>3.06**  5.49</td>
<td>0.36  0.99</td>
<td>1.67  4.16</td>
<td>0.62  1.60</td>
</tr>
<tr>
<td>Late response</td>
<td>0.36  0.93</td>
<td>0.45  0.84</td>
<td>0.44  1.12</td>
<td>2.55***  2.66</td>
<td>0.30  0.68</td>
<td>0.74  1.41</td>
<td>0.26  0.50</td>
</tr>
<tr>
<td>Misses</td>
<td>0.90  2.19</td>
<td>3.77**  7.05</td>
<td>1.01  2.65</td>
<td>2.79*  3.66</td>
<td>0.88  1.22</td>
<td>3.95*  6.98</td>
<td>0.67  0.98</td>
</tr>
</tbody>
</table>

* p<0.05,  ** p<0.01,  *** p<0.001  compared to respective control comparison data

a Each control comparison was resampled from the overall control group (n=138) so are not independent (see methods).

b For n=4 datasets (1.3%), sigma was returned as 0 in the ex-Gaussian model.

c Mann-Whitney U test.
**Title:** The Vigil Continuous Performance Test

**Legend:** Stimulus timing and example of a re-classification of a late response to a target sequence (left).

General response classification rules (right). *If the previous stimulus was a target, the algorithm first checked if this target already had a valid response, in which case the current response was also classified as a commission error. This path is omitted in the figure.*
Figure 2.

*Title:* Vincentile plots for all clinical groups compared to matched control data

**Legend:** V1 to V8 denotes each Vincentile (sequential 12.5% of RT data) from fastest to slowest RT.