Psychological Medicine

www.cambridge.org/psm

Characterizing intraindividual variability in bipolar disorder: links to cognition, white matter microstructure, and clinical variables

Georgia F. Caruana[1], Sean P. Carruthers[2], James A. Karantonis[1,2], Lisa S. Furlong[1], Eric J. Tan[3,4,5], Erica Neill[6,7], Susan L. Rossell[2,5] and Tamsyn E. Van Rheenen[1,2]

[1]Department of Psychiatry, The University of Melbourne, Melbourne, VIC, Australia; [2]Centre for Mental Health, School of Health Sciences, Swinburne University, Melbourne, VIC, Australia; [3]Memory, Aging and Cognition Centre, National University Health System, Singapore; [4]Department of Pharmacology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore; [5]St Vincent's Mental Health, St Vincent's Hospital, Melbourne, VIC, Australia; [6]Orygen, Centre for Youth Mental Health, The University of Melbourne, Melbourne, VIC, Australia and [7]Centre for Youth Mental Health, The University of Melbourne, Melbourne, VIC, Australia

Original Article

Cite this article: Caruana, G. F., Carruthers, S. P., Karantonis, J. A., Furlong, L. S., Tan, E. J., Neill, E., Rossell, S. L., & Van Rheenen, T. E. (2025). Characterizing intraindividual variability in bipolar disorder: links to cognition, white matter microstructure, and clinical variables. *Psychological Medicine*, **55**, e204, 1–10
https://doi.org/10.1017/S0033291725100731

Received: 09 February 2025
Revised: 13 May 2025
Accepted: 19 May 2025

Keywords:
attention; bipolar disorder; coefficient of variation; cognition; fractional anisotropy; heterogeneity; individual standard deviation; intraindividual variability; neuropsychology; neurocognition; radial diffusivity; white matter microstructure

Corresponding author:
Georgia F. Caruana and Tamsyn E Van Rheenen;
Emails: gcaruana@student.unimelb.edu.au; tamsyn.van@unimelb.edu.au

Abstract

Background. Most cognitive studies of bipolar disorder (BD) have examined case–control differences on cognitive tests using measures of central tendency, which do not consider intraindividual variability (IIV); a distinct cognitive construct that reliably indexes meaningful cognitive differences between individuals. In this study, we sought to characterize IIV in BD by examining whether it differs from healthy controls (HCs) and is associated with other cognitive measures, clinical variables, and white matter microstructure.

Methods. Two hundred and seventeen adults, including 100 BD outpatients and 117 HCs, completed processing speed, sustained attention, working memory, and executive function tasks. A subsample of 55 BD participants underwent diffusion tensor imaging. IIV was operationalized as the individual standard deviation in reaction time on the Continuous Performance Test-Identical Pairs version.

Results. BD participants had significantly increased IIV compared to age-matched controls. Increased IIV was associated with poorer mean performance scores on processing speed, sustained attention, working memory, and executive function tasks, as well as two whole-brain white matter indices: fractional anisotropy and radial diffusivity.

Conclusions. IIV is increased in BD and appears to correlate with other cognitive variables, as well as white matter measures that index reduced structural integrity and demyelination. Thus, IIV may represent a neurobiologically informative cognitive measure for BD research that is worthy of further investigation.

Introduction

Cognitive impairment is a common and often debilitating feature of bipolar disorder (BD) that persists across different mood states and confers substantial functional and psychosocial burden (Bora, Yucel, & Pantelis, 2009; Bortolato et al., 2015; Burdick et al., 2014; Cullen et al., 2016; Karantonis et al., 2020; Simonsen et al., 2010). Most cognitive research on BD has analyzed case–control differences in measures of central tendency, which, as recently critiqued by Sánchez-Torres et al. (2023), can mask individual fluctuations and variations in cognitive performance that are clinically relevant in psychiatry. Indeed, this within-person variation in cognitive performance, otherwise known as *intraindividual variability* (IIV) (MacDonald, Hultsch, & Dixon, 2003), maybe a feature of cognition in BD.

IIV can be measured across tasks and/or time but is most typically operationalized as trial-to-trial response consistency *within* a single reaction time task using the metrics of *individual standard deviation* (iSD) and/or the *individual coefficient of variation* (CoV) (Christensen et al., 2005; Hultsch & MacDonald, 2004). The former reflects the iSD of item-by-item response times. The latter reflects the ratio of the iSD of response times to the mean of those response times and is calculated by dividing the iSD by the individual mean. Increases in these IIV measures indicate irregular cognitive performance, potentially mediated by abnormalities in top-down executive control (MacDonald, Hultsch, & Dixon, 2003).

Originally explored in the context of aging, IIV was first considered to reflect psychometric noise, but is now recognized as a distinct cognitive construct that reliably indexes meaningful cognitive differences between individuals (MacDonald, Hultsch, & Dixon, 2003; Ram, Rabbitt, Stollery, & Nesselroade, 2005). Increased IIV has been associated with poorer socio-occupational functioning (Fuermaier et al., 2015; Rajji, Miranda, & Mulsant, 2014) and quality of life (Mitchell,

© The Author(s), 2025. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives licence (http://creativecommons.org/licenses/by-nc-nd/4.0), which permits non-commercial re-use, distribution, and reproduction in any medium, provided that no alterations are made and the original article is properly cited. The written permission of Cambridge University Press must be obtained prior to any commercial use and/or adaptation of the article.

Kemp, Benito-León, & Reuber, 2010), as well as increased rates of mortality (Deary & Der, 2005). IIV increases have also been observed in healthy elderly people (Christensen et al., 2005; Hultsch, MacDonald, & Dixon, 2002; MacDonald, Li, & Bäckman, 2009) and found to predict executive function, memory, processing speed, and sustained attention impairments within them several years later (Bielak et al., 2010a, 2010b; Cherbuin, Sachdev, & Anstey, 2010; MacDonald, Hultsch, & Dixon, 2003). Increased IIV (abnormal to that expected by age) is also a marker of the Alzheimer's disease prodrome (Roalf et al., 2018), and has been found to index early cognitive changes in people with other neurodegenerative conditions who are otherwise not yet demonstrating cognitive impairments (Jones, Burroughs, Apodaca, & Bunch, 2020; Kälin et al., 2014; Mazerolle, Wojtowicz, Omisade, & Fisk, 2013; Wojtowicz, Omisade, & Fisk, 2013). This may relate to the demonstrated correlations of IIV with the brain's white matter (MacDonald, Nyberg, & Bäckman, 2006; Nilsson, Thomas, O'Brien, & Gallagher, 2014), changes in which have been found to precede the onset of observable cognitive impairments by several years (Silbert et al., 2012). Indeed, several studies have found that increases in IIV are associated with reduced white matter volume (Jackson, Balota, Duchek, & Head, 2012; Lövdén et al., 2013) and microstructural integrity in major frontal, parietal, and central white matter networks (Fjell, Westlye, Amlien, & Walhovd, 2011; Halliday, Gawryluk, Garcia-Barrera, & MacDonald, 2019; Mella, de Ribaupierre, Eagleson, & de Ribaupierre, 2013; Moy et al., 2011; Tamnes et al., 2012).

It has been argued that IIV may index clinically and biologically meaningful information better than measures of central tendency alone (Davis, Sivaramakrishnan, Rolin, & Subramanian, 2025; Dykiert, Der, Starr, & Deary, 2012; MacDonald, Hultsch, & Dixon, 2003; Sánchez-Torres et al., 2023; Tamnes et al., 2012; Williams et al., 2005). Hence, examining IIV in BD could expand our insights into the cognitive profile of the disorder and help to elucidate the putative mechanisms contributing to the associated cognitive impairments, which, to date, remain unknown. Only a handful of preliminary studies have examined IIV in BD, finding it to be increased, on average, in middle-aged adults (Gallagher et al., 2015; Haatveit et al., 2023; Krukow et al., 2017) and youth samples (Brotman et al., 2009) compared to age-matched controls. One study showed that increased IIV in BD persisted longitudinally and was negatively associated with a global index of cognition (Depp et al., 2012), while another study found that IIV increased even further as a function of the complexity of the cognitive task used (Moss et al., 2016). Only a few studies have examined whether IIV in BD is associated with clinical variables, presenting mixed findings regarding the role of mood symptoms, medication load, age of onset, or illness duration (Depp et al., 2012; Gallagher et al., 2015; Haatveit et al., 2023; Krukow et al., 2017; Moss et al., 2016). Moreover, no studies have examined how IIV in BD relates to other cognitive domains or to indices of white matter, abnormalities of which are observed in BD and linked, to some extent, to its cognitive symptoms (Caruana et al., 2024).

Considering the above, the characterization of IIV in BD remains in its infancy. The replication and expansion of existing preliminary studies focused on IIV is thus required to determine the extent to which IIV can inform our broader understanding of cognition within the disorder. In this study, we aimed to do this by further characterizing IIV and its correlates with BD. We specifically sought to replicate prior findings showing increased IIV in patients with BD compared to controls using a larger sample than most previous research. We also aimed to determine whether IIV (i) is particularly related to any specific cognitive domain, (ii) covaries with clinical symptoms, and (iii) is associated with whole-brain white matter integrity. We hypothesized that IIV would be increased in people with BD compared to controls, and that this increased IIV in BD would be associated with poorer cognitive performance across a range of domains as well as decreased whole-brain white matter integrity. The extent of associations between IIV and clinical variables remained an open question.

Methods

The study was approved by the local Human Ethics Review Committee and adhered to the Declaration of Helsinki.

Participant characterization

The data from 217 participants (n = 100 with BD and n = 117 healthy controls [HCs]) were included in this study. All participants had participated in studies led by the authors (e.g. see Karantonis et al., 2020; Neill & Rossell, 2013; Tan & Rossell, 2014; Van Rheenen & Rossell, 2014b) and had been recruited using general advertisements as well as online websites and social media, with the BD participants also being recruited through community support groups. All participants had given prior informed consent for the analysis of their data.

Participants were aged between 18 and 65 years, were proficient in English, and had no known neurological disorders, acute medical illnesses, or significant hearing or visual impairments, no current alcohol or substance abuse/dependence, and none were pregnant. HCs also had no first-degree relatives with a psychiatric diagnosis. BD diagnosis and HC eligibility were confirmed using the Mini-International Neuropsychiatric Interview for BD (Sheehan et al., 1998), with 83 BD participants meeting criteria for BD-I and 17 meeting criteria for BD-II. BD participants were all clinically stable outpatients at the time of assessment, and none were experiencing symptoms of psychosis. Current mood symptom severity was measured using the Montgomery and Åsberg Depression Rating Scale (MADRS; Montgomery & Åsberg, 1979) and the Young Mania Rating Scale (YMRS; Young, Biggs, Ziegler, & Meyer, 1978). Sixty-six percent of participants were considered effectively stable with MADRS scores <12 and YMRS <8. The remaining 34% were symptomatic with mild–moderate symptoms (76% of these had MADRS scores >12, and a further 24% had YMRS scores >8). Self-reported use of mood stabilizers, antipsychotics, and antidepressants in the sample was also recorded, as was the age of illness onset, illness duration, psychiatric hospitalizations, and mood episode history (Table 1).

Intraindividual variability

IIV measures were derived for all participants from individual responses on the Continuous Performance Test-Identical Pairs (CPT-IP) version (Cornblatt et al., 1988). The CPT-IP was collected during the administration of the Matrics Consensus Cognitive Battery (MCCB), a battery of tests validated for use in BD (Burdick et al., 2011; Van Rheenen & Rossell, 2014a). The CPT-IP is a computerized neurocognitive measure requiring participants to monitor a series of two-, three-, and then four-digit sequences and respond when identical sequences are presented consecutively. Across the two-, three-, and four-digit blocks, a total of 450 rapidly flashed digit sequences (150 per block) are delivered, including 30 'target' digit pairs within each block, as well as 30 'catch' trials that feature two successive similar but not identical digit sequences, and 90 random digit sequences that are in no way similar. Stimuli are flashed on the screen for 50 ms, followed by a 950-ms blank screen (stimulus onset asynchrony = 1,000 ms). Participants are asked to respond by quickly pressing and

https://doi.org/10.1017/S0033291725100731 Published online by Cambridge University Press

Table 1. Demographic, clinical, and cognitive characteristics of the full sample

	Bipolar disorder (BD), *n* = 100			Healthy controls (HC), *n* = 117			Group comparisons			
	n (%)	M	SD	*n* (%)	M	SD	t/χ^2	*df*	*p*	Cohen's *d*
Demographic										
Age (years)		38.382	12.090		36.664	13.357	0.989	215	0.324	0.135
Sex (m/f)	49(49.0)/51(51.0)			49(41.88)/68(58.12)			1.104	1	0.293	0.071
Education (years)[+]		16.341	4.122		16.520	3.602	−0.335	204	0.738	0.046
Estimated premorbid IQ[^]		109.344	11.242		110.732	9.013	−1.003	214	0.317	0.136
Clinical										
Age of onset (years)		22.207	9.976							
Illness duration (years)		13.374	10.189							
Lifetime number of mood episodes (self-reported)[#]										
Mania		22.263	41.916							
Depression		32.874	60.879							
History of psychosis	58 (58.0)									
Lifetime number of hospitalizations (self-reported)		3.080	5.675							
MADRS		9.422	8.861							
YMRS		4.174	4.426							
Pharmacotherapy (% using)										
Mood stabilizer	62 (62.0)									
Antipsychotic	41 (41.0)									
Antidepressant	31 (31.0)									
No medication	12 (12.0)									
Total medication load		2.230	1.514							
Cognitive										
Global IIV measures										
iSD		120.788	29.851		106.001	25.081	15.724	1,215	≤0.001*	0.536
CoV		0.221	0.045		0.202	0.041	10.830	1,215	0.001*	0.442
Sustained attention										
CPT-IP – average d-prime		2.634	0.672		2.902	0.515	−3.306	213	0.001*	0.448
Executive function										
Trail-making test – B[~!]		62.615	38.233		50.508	19.038	−1.839	–	0.066	0.401
Working memory										
Letter number span[~]		15.470	3.151		16.393	2.681	−1.973	–	0.049	0.316
Processing speed										
BACS-symbol coding		55.110	10.920		62.853	11.432	−5.067	214	<0.001*	0.692

Note: CoV, coefficient of variation; IIV, intraindividual variability; iSD, individual standard deviation; MADRS, Montgomery and Åsberg Depression Rating Scale; YMRS, Young Mania Rating Scale; + total years of education missing from five BD and seven HC participants; ^ premorbid IQ score missing from two BD participants; # number of mood episodes missing from 10 BD participants; *p*-values reported in the table reflect raw values, but are designated with a * if remaining significant after FDR correction; ~ Mann–Whitney *U*-test; ! higher trail-making test-B scores indicate worse performance.

releasing the left mouse button whenever they are presented with an identical digit sequence pair. The total test time is ~10 min.

IIV was operationalized from the raw reaction time (milliseconds) values for each successful 'hit' across the entirety of the CPT-IP. (The size of the case–control difference in the mean standard deviation in response time for the CPT-IP was stronger when calculated across its entirety versus for each individual block. Hence, we decided to use the IIV measures calculated across the full task. See the Supplementary Methods and Supplementary Table S1 for further details.) Hit trials refer to any trial in which the second stimulus in a target pair received a response during the interstimulus interval (i.e. the response window). Thus, any response to the second

https://doi.org/10.1017/S0033291725100731 Published online by Cambridge University Press

stimulus in a target pair occurring after the onset of the next trial was not retained for the IIV calculation. Each participant's mean reaction time was based on their hits, and the standard deviation of the mean of this reaction time (reflecting the participant's *iSD*) was computed. Each individual's *CoV* was also derived by dividing the *iSD* by their mean reaction time. These IIV measures are mathematically (and conceptually) distinct from accuracy, speed, or d-prime scores typically derived from the CPT-IP (Cho et al., 2023).

Other cognitive measures

Premorbid IQ was estimated in all participants using the Wechsler Test of Adult Reading (Wechsler, 2001). Scores from other relevant and available cognitive tests were also analyzed. These included overall mean *d-prime* scores, which are the standard metrics of *sustained attention* from the CPT-IP in that they reflect a ratio of speed, accuracy, and focus when discriminating between target and distractor digit sequences across blocks. Scores from the Brief Assessment of Cognition in Schizophrenia-Symbol Coding were used as a measure of *processing speed*, and from the Letter Number Span as a measure of *working memory*. Time to complete scores from the Trail Making Test: Part B (Reitan, 1958) were used as a measure of *executive function*. All subtests are described in detail elsewhere (Kern et al., 2008; Nuechterlein et al., 2008; Yatham et al., 2010). Better performance was represented by higher scores on the working memory, processing speed, and sustained attention measures and lower scores on the executive measure.

Neuroimaging acquisition and processing

Diffusion-weighted imaging (DWI) scans were available in a *subset of the BD sample* ($n = 55$, of which $n = 52$ had BD I, $n = 3$ had BD II; 33 males and 28 females). Scans were acquired on a Siemens Magnetom 3T Tim Trio system (Erlangen, Germany) using a 34-channel head coil and a multi-shell protocol (Repetition Time = 9200 ms, Echo Time = 117 ms, voxel size = $2.5 \times 2.5 \times 2.5$ mm^3). Sixty diffusion gradient directions were acquired with a *b*-value of 3,000 s/mm^2, 30 directions with a *b*-value of 2,000 s/mm^2, and 30 directions with a *b*-value of 900 s/mm^2. Ten non-diffusion-weighted images (b-value = 0 s/mm^2) were also acquired, and to enable the estimation of susceptibility-induced off-resonance fields, five additional non-diffusion-weighted images were acquired with the same imaging parameters but with a reversed-phase encoding direction.

Images were processed and analyzed using FMRIB Software Library 6.0.1 (Smith et al., 2004), adhering to the ENIGMA-DWI protocol. Susceptibility-induced distortions were estimated and corrected for using the TOPUP method (Andersson, Skare, & Ashburner, 2003). Subject motion and eddy current-induced distortion correction, as well as automated outlier replacement, were performed in line with the methods (EDDY) described in Andersson and Sotiropoulos (2016) and Andersson, Graham, Zsoldos, and Sotiropoulos (2016). Fractional anisotropy (FA) maps were estimated for each participant using the DTIFIT option with the FMRIB Diffusion Toolbox by fitting a tensor model to the pre-processed diffusion data. Axial diffusivity (AD) ($\lambda 1$) and radial diffusivity (RD) ($[\lambda 2 + \lambda 3]/2$) maps were also estimated using the eigenvalues associated with the fitted tensor model. Using tract-based spatial statistics, participant FA maps were then aligned to the custom ENIGMA-DWI FA template derived from 400 adult participants (Jahanshad et al., 2013) and subsequently projected onto the ENGIMA-DWI template skeleton. The same method was used to project images of FA's constituent measures: mean diffusivity (MD), AD, and RD onto the skeleton. Voxels along the individual skeletons were averaged across 25 bilateral regions of interest (ROIs) based on the JHU WM atlas (Mori et al., 2008). Each of the diffusion measures was then imported into the Statistical Package for the Social Sciences (SPSS) and averaged over all ROIs to generate whole-brain FA, MD, AD, and RD values for each participant.

Statistical analyses

All data were statistically analyzed using SPSS version 27 (IBM). First, variables were visually checked for extreme outliers, and relevant statistical test assumptions were assessed and met using standard methods. (Outliers were considered at the sample level after iSD and CoV for each participant had been calculated, based on the SPSS categorizations of extreme outliers; that is, iSD and CoV scores that were less than/greater than three SDs of the mean were excluded.) In preliminary analyses, two-tailed comparisons (using χ^2-tests and independent-sample *t*-tests as appropriate) were conducted to compare the BD and HC groups on relevant demographic and clinical variables to characterize abnormalities in the BD sample and identify any potential covariates. The association of these potential covariates and the IIV indices (iSD and CoV) was also examined in the full sample using Pearson's correlations, with only the variables that were significantly correlated with IIV being covaried in the analyses.

In the primary analyses, group differences in mean *iSD*, *CoV*, and other cognitive test scores were ascertained using one-way analyses of variance (ANOVAs). Two-tailed bivariate Pearson's correlations (or nonparametric equivalent tests) were then conducted to explore associations between the IIV indices and the other cognitive scores in the BD and HC groups separately, with Fisher's *Z*-tests used to compare correlations. Correlations were also conducted in the BD imaging subsample to examine the associations between whole-brain white matter (FA, RD, MD, and AD) and global IIV indices, as well as other cognitive test scores. Furthermore, associations between IIV and mood symptom severity scores were examined in the BD sample and in relation to other clinical factors (age of onset, illness duration, number of mood episodes, hospitalizations due to mood disturbance, and medication load). Group comparisons in mean IIV based on BD diagnostic subtype, psychosis history, and usage of key medication types (mood stabilizers, antipsychotics, and antidepressants) were also conducted using ANOVA (or nonparametric equivalent). To examine the influence of mood state and establish state versus trait effects, group comparisons of mean IIV were conducted, excluding the symptomatic BD participants (using the same procedures as used in the full sample) in secondary sensitivity analyses.

A false discovery rate (FDR) of $p < 0.05$ was applied to the results to account for multiple comparisons using the Benjamini–Hochberg method (see the Supplementary Material for details). The effect sizes in the text are given in Cohen's *d*, while the reported *p*-values reflect raw, uncorrected values.

Results

Demographic, clinical, and cognitive characteristics of the sample

Demographic, clinical, and cognitive characteristics of the full sample are provided in Table 1 and of the DWI subsample in Supplementary Table S2. There were no differences in age, sex,

https://doi.org/10.1017/S0033291725100731 Published online by Cambridge University Press

years of education, or estimated premorbid IQ between the BD and HC groups, and none of these characteristics were significantly associated with *iSD* (age r = .098, p = .150; years of education r = −.040, p = .566; estimated premorbid IQ r = −.127, p = 0.063) or *CoV* (age r = .017, p = .799; years of education r = −.074, p = .290; estimated premorbid IQ r = −.057, p = .401). The mean MADRS and YMRS scores in the BD sample were low (M = 9.422, SD = 8.861 and M = 4.174, SD = 4.26, respectively). Fifty-eight percent of the BD group had a *history* of psychosis. Cognitively, the BD group performed significantly more poorly than HCs on measures of sustained attention and processing speed, with moderate effects (d = 0.448 and d = 0.692, respectively). The BD group also had worse mean executive function and working memory than HCs, with small effects (d = 0.401 and 0.316), although the former comparison was not significant initially and the latter did not survive the FDR correction. Moreover, BD participants had significantly higher mean IIV (indexed by both iSD and CoV) than HCs (Figure 1; iSD [F(1,215) = 15.724, $p \leq 0.001$] and CoV [F(1,215) = 10.830, p = 0.001]).

Associations between global IIV indices and other cognitive tests in BD and HC groups

In the BD group, a higher mean IIV, as measured by both *iSD* and *CoV*, was associated with worse performance in all four cognitive domains analyzed: sustained attention, working memory, executive function, and processing speed (Figure 2a). In the HC group, a higher mean IIV was only associated with worse sustained attention (Figure 2b). A secondary check of these associations using median splitting within each group based on high/low IIV indicated that these results persisted (Supplementary Table S3); however, no Fisher's Z comparisons of the correlations between groups were significant.

Associations between global IIV indices and clinical variables in the full BD sample

Mean iSD was higher in BD participants using antidepressants (on antidepressants: M = 129.949, SD = 32.573; off antidepressants: M = 115.958, SD = 28.854; F(1,86) = 4.277, p = 0.042, d = 0.455). Mean CoV was higher in BD-I than BD-II participants (BD-I: M = 0.225, SD = 0.047; BD-II: M = 0.202, SD = 0.028; Kruskal–Wallis test p = 0.025, d = 0.595) and in BD participants using versus not using antipsychotics (on antipsychotics: M = 0.233, SD = 0.047; off antipsychotics: M = 0.213, SD = 0.045; F(1,84) = 3.956, p = 0.050, d = 0.435). Mean CoV was also negatively correlated with age of BD onset (r = −0.258, p = 0.016). However, none of these results survived FDR correction. No other significant associations were observed. See Supplementary Tables S4 and S5 for details.

Sensitivity analyses including only affectively stable BD participants (n = 66)

There were no substantial differences between the outcomes of the sensitivity analyses that *excluded symptomatic BD participants* and the analyses conducted using the full BD sample (Supplementary Tables S6 and S7).

Associations between IIV indices and global white matter microstructure in the BD imaging subset (n = 55)

Figure 3 shows a significant negative association between whole-brain FA and mean *iSD* (r = −.275, p = .042). Significant positive associations were also evident between whole-brain RD and mean *iSD* (r = .271, p = .045) and mean *CoV* (r = .296, p = .028). No associations between the IIV indices and whole-brain MD or AD

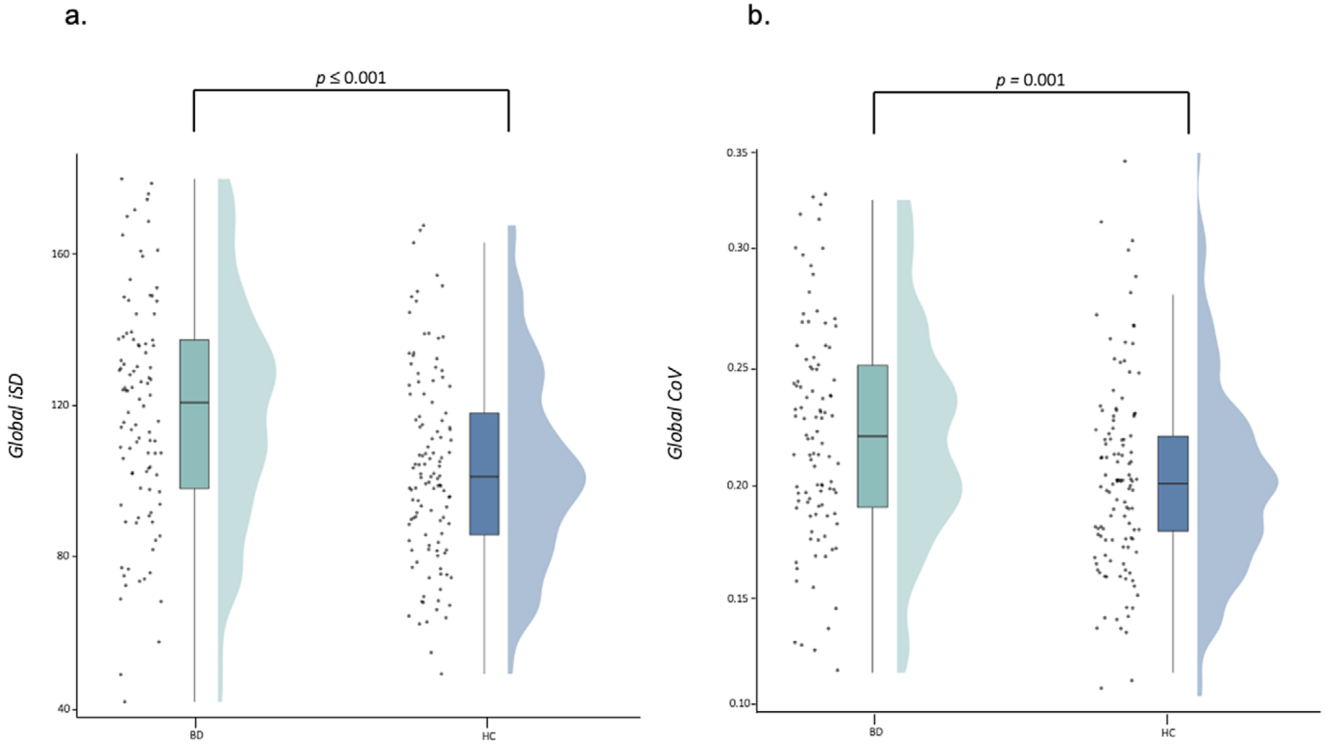


Figure 1. Raincloud plots depicting mean comparisons of (a) global iSD and (b) global CoV between bipolar disorder (BD) and healthy control (HC) groups. *p*-Values reflect raw values, but are significant after FDR correction. CoV, 'coefficient of variation'; iSD, 'individual standard deviation'.

https://doi.org/10.1017/S0033291725100731 Published online by Cambridge University Press

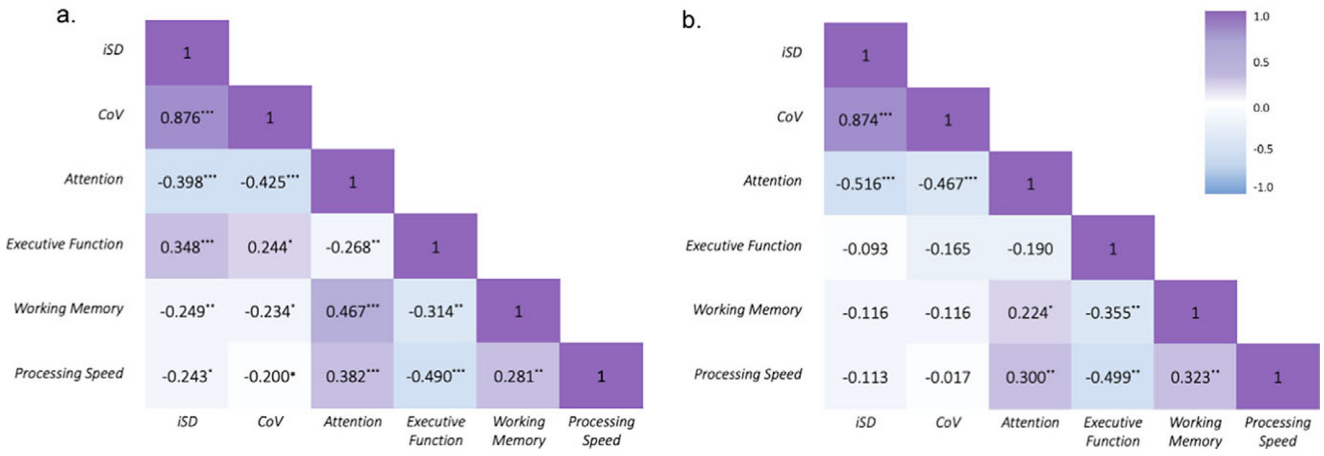


Figure 2. Spearman's rho correlations between IIV indices and the different cognitive domains for the (a) bipolar disorder (BD) and (b) healthy control (HC) groups.
Note: CoV, 'coefficient of variation'; iSD, 'individual standard deviation'; * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$ (FDR-corrected).

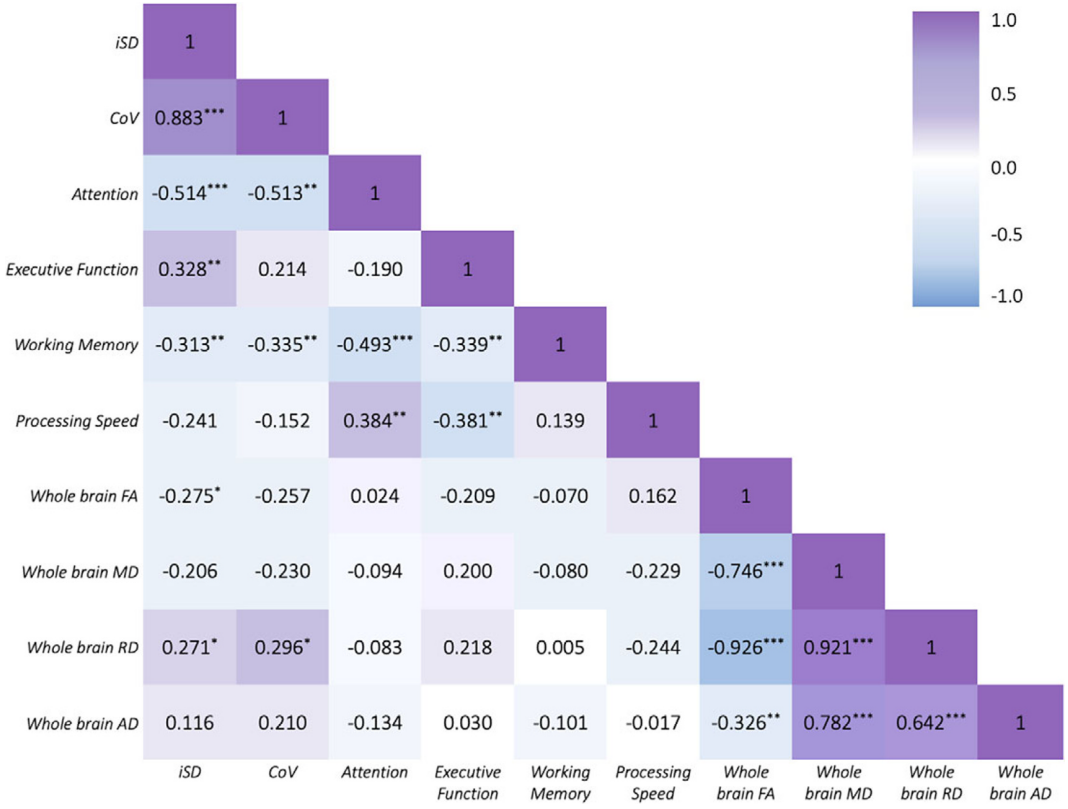


Figure 3. Pearson's *r* correlations of global IIV indices with diffusion-weighted imaging measures and the different cognitive domain scores in the *BD neuroimaging subsample.*
Note: AD, 'axial diffusivity'; CoV, 'coefficient of variation'; FA, 'fractional anisotropy'; iSD, 'individual standard deviation'; MD, 'mean diffusivity'; RD, 'radial diffusivity'; *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$ (FDR-corrected).

were evident, nor were there associations between any of the DWI measures and sustained attention, executive function, working memory, or processing speed.

Discussion

This study expanded the characterization of IIV in BD, identifying mean increases in IIV in BD as compared to HCs of an equivalent mean age and sex. This increased IIV in BD aligns with our hypothesis and replicates and reinforces the results of prior, albeit smaller, IIV BD studies (Brotman et al., 2009; Depp et al., 2012; Gallagher et al., 2015; Haatveit et al., 2023; Krukow et al., 2017; Moss et al., 2016). That the average IIV of the BD sample remained increased in sensitivity analyses that removed all symptomatic patients, supports suggestions that IIV abnormalities in BD are trait-like in nature (Depp et al., 2012). Moreover, these increases were not linked to clinical variables, such as a longer illness duration, a history of more mood episodes, or psychiatric hospitalizations.

https://doi.org/10.1017/S0033291725100731 Published online by Cambridge University Press

IIV increases in BD were found in individuals with worse processing speed and working memory performance, as well as in those with poorer sustained attention and executive function. The largest effects were for the latter two cognitive domains, consistent with evidence that IIV closely covaries with, and may even proxy, top-down attentional and executive control (Cañigueral et al., 2023; MacDonald, Li, & Bäckman, 2009). In the HC group, IIV was significantly associated with only sustained attention, which is sensical, given that sustained attention scores were derived from the same cognitive task as the IIV measures. Thus, a correlation between them is expected even though these measures are conceptually and mathematically distinct. The significant correlations between IIV and the other cognitive domain scores were observed solely in the BD group, which may reflect that increased IIV in this group is one index of a more generalized cognitive impairment. This is consistent with theories that domain-level cognitive abilities are more related to each other (less differentiated) at lower levels of general cognitive ability than they are at higher levels (Tucker-Drob, 2009). This inference should be interpreted with the caveat that the differences in correlations between the BD and HC groups were nonsignificant.

In our study, increased IIV in BD was associated with two measures of white matter integrity: reduced whole-brain FA and increased whole-brain RD. Since concurrent decreases in FA and increases in RD may reflect damaged white matter resulting from reduced myelin integrity/demyelination (Johnson, Diaz, & Madden, 2015; Madden, Bennett, & Song, 2009), this pattern of findings aligns with previous work positing that one neurobiological mechanism underpinning increases in IIV is reduced action potential conduction efficiency caused by axonal or myelin abnormalities (Fjell, Westlye, Amlien, & Walhovd, 2011; Moy et al., 2011). It is notable that in our study, FA and RD were only correlated with the IIV metrics of interest and not with any other cognitive domain scores. Thus, IIV appears to provide unique information about the brain-behavior relationship in BD, beyond that of the more commonly used cognitive scores, which have not been robustly linked to white matter microstructure in the disorder to date (Caruana et al., 2024).

Taken together, our data demonstrate that increases of IIV in BD that are associated with reductions in performance in other cognitive domains, as well as in white matter microstructural integrity – patterns typically observed in aging samples (Nilsson, Thomas, O'Brien, & Gallagher, 2014). In our data, increases in IIV in BD were evident, despite the BD and HC groups being equated in terms of mean age, and age not being significantly associated with IIV in the BD group or the overall sample. Given that, increased IIV may be considered a marker of advancing age, and considering that this sample largely comprised middle-aged adults within a period of the lifespan generally characterized by cognitive consistency and stability (Ferreira et al., 2017), we speculate that the increased IIV in BD observed here may reflect the outcome of premature or accelerated cognitive aging. An alternative explanation is that elevated IIV in BD is related to a lag in normative cognitive development, given that IIV is known to follow a U-shape curve across the lifespan in which it is initially high during childhood, plateaus during adulthood, and trends upward in the elderly (MacDonald, Li, & Bäckman, 2009). However, since cognition is not typically impaired in BD during the premorbid period (which typically coincides with childhood and adolescence (Van Rheenen et al., 2020), this explanation seems less likely.

Some limitations of this study should be considered. First, IIV metrics were calculated from individual responses to target stimuli that occurred within a 949-ms response window. This would have disadvantaged particularly slow participants whose responses outside this period would not have been captured. Second, the use of cross-sectional data precluded our ability to test the directionality of relationships between IIV and other variables of interest. Third, the imaging subsample was modest in size and comprised only those with BD, limiting our ability to conduct white matter tract-specific analyses or group comparisons with HCs. Fourth, the effects in this study were small to moderate in size, which should be taken into consideration when interpreting the results. That said, the absence of large effects suggests that other factors of relevance to IIV may be relevant to future research on this topic, such as peripheral inflammation, stress, and trauma, which are implicated in BD *and* known correlates of cognitive performance and white matter pathology (Li, Xu, & Wang, 2023).

Finally, it should be mentioned that our use of the CPT-IP to examine IIV was based on the availability of this test within a widely used cognitive battery, the MCCB. iSD and CoV are easily calculated from the CPT-IP and were thus used to operationalize IIV here. However, these IIV indices assume that the response times across the CPT-IP are Gaussian (i.e. normally distributed) (Moss et al., 2016). We initially reasoned that the use of such metrics was most suitable because ex-Gaussian analyses require ~100 trials as a rule of thumb, and the absolute number of trials from which valid responses can be recorded from the CPT-IP for each digit-sequence block is limited to 30. However, subsequent preliminary analyses in our data suggested that IIV is best measured across the totality of the CPT-IP, since group differences in IIV were found not to be affected by the increasing cognitive load of each digit block, and the largest effects were evident when all valid responses across the task were used to calculate the IIV metrics (see Supplementary Materials for details). This suggests that the CPT-IP *may* be suitable for ex-Gaussian analyses, as there are 90 possible hit trials across all blocks. Thus, future extensions of our work could benefit from examining ex-Gaussian parameters, such as mu, sigma, and tau, in addition to iSD and CoV.

In summary, this study complements an existing, albeit small, evidence base showing that IIV is increased in BD. It extends it by demonstrating that IIV elevations can be elicited from a widely used cognitive test from the MCCB using easily calculated metrics that are detrimentally associated with cognitive performance across other domains, as well as with a proxy of underlying myelin damage in the neural white matter. Given the unique links between IIV and white matter, but not between white matter and more traditionally used cognitive scores, IIV may be considered a neurobiologically informative cognitive measure for BD that is worthy of future research.

Supplementary material. The supplementary material for this article can be found at http://doi.org/10.1017/S0033291725100731.

Acknowledgments. The authors would like to acknowledge the facilities and scientific assistance of the National Imaging Facility, a National Collaborative Research Infrastructure Strategy (NCRIS) capability, at the Swinburne Neuroimaging Facility, Swinburne University of Technology. The authors would also like to sincerely thank the participants in this study for their time and contributions. The authors note that for the purposes of open access, a CC BY public copyright licence has been applied to any Author Accepted Manuscript version arising from this submission.

Funding statement. In completing this work, TVR was supported by the University of Melbourne and an Al and Val Rosenstrauss Fellowship from the Rebecca L Cooper Medical Research Foundation. SLR was supported by National Health and Medical Research Council (NHMRC) Senior Fellowship (GNT1154651) and project grant (GNT1060664). GFC was supported by the

University of Melbourne/Fay Marles Bursary. JK was supported by a Swinburne University/postgraduate scholarship and LF was supported by an Australian Rotary Health/Ian Parker Bipolar Research Fund/Brunslea Park Estate postgraduate scholarship. Data collection was supported by the Jack Brockhoff Foundation, University of Melbourne, Barbara Dicker Brain Sciences Foundation, Rebecca L Cooper Foundation, the Society of Mental Health Research, Helen McPherson Smith Trust and Australian Rotary Health.

Competing interests. The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflicts of interest.

References

Andersson, J. L., Skare, S., & Ashburner, J. (2003). How to correct susceptibility distortions in spin-echo echo-planar images: Application to diffusion tensor imaging. *NeuroImage*, **20**(2), 870–888. https://doi.org/10.1016/s1053-8119(03)00336-7.

Andersson, J. L. R., Graham, M. S., Zsoldos, E., & Sotiropoulos, S. N. (2016). Incorporating outlier detection and replacement into a non-parametric framework for movement and distortion correction of diffusion MR images. *NeuroImage*, **141**, 556–572. https://doi.org/10.1016/j.neuroimage.2016.06.058.

Andersson, J. L. R., & Sotiropoulos, S. N. (2016). An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *NeuroImage*, **125**, 1063–1078. https://doi.org/10.1016/j.neuroimage.2015.10.019.

Bielak, A. A., Hultsch, D. F., Strauss, E., MacDonald, S. W., & Hunter, M. A. (2010a). Intraindividual variability in reaction time predicts cognitive outcomes 5 years later. *Neuropsychology*, **24**(6), 731.

Bielak, A. A., Hultsch, D. F., Strauss, E., MacDonald, S. W., & Hunter, M. A. (2010b). Intraindividual variability is related to cognitive change in older adults: Evidence for within-person coupling. *Psychology and Aging*, **25**(3), 575.

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. *Journal of Affective Disorders*, **113**(1–2), 1–20. https://doi.org/10.1016/j.jad.2008.06.009.

Bortolato, B., Miskowiak, K. W., Köhler, C. A., Vieta, E., & Carvalho, A. F. (2015). Cognitive dysfunction in bipolar disorder and schizophrenia: A systematic review of meta-analyses. *Neuropsychiatric Disease and Treatment*, **11**, 3111–3125. https://doi.org/10.2147/ndt.S76700.

Brotman, M. A., Rooney, M. H., Skup, M., Pine, D. S., & Leibenluft, E. (2009). Increased intrasubject variability in response time in youths with bipolar disorder and at-risk family members. *Journal of the American Academy of Child and Adolescent Psychiatry*, **48**(6), 628–635. https://doi.org/10.1097/CHI.0b013e3181a27527.

Burdick, K. E., Goldberg, T. E., Cornblatt, B. A., Keefe, R. S., Gopin, C. B., DeRosse, P., & Malhotra, A. K. (2011). The MATRICS consensus cognitive battery in patients with bipolar I disorder. *Neuropsychopharmacology*, **36**(8), 1587–1592. https://doi.org/10.1038/npp.2011.36.

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. *Psychological Medicine*, **44**(14), 3083–3096. https://doi.org/10.1017/s0033291714000439.

Cañigueral, R., Ganesan, K., Smid, C. R., Thompson, A., Dosenbach, N. U. F., & Steinbeis, N. (2023). Intra-individual variability adaptively increases following inhibition training during middle childhood. *Cognition*, **239**, 105548. https://doi.org/10.1016/j.cognition.2023.105548.

Caruana, G. F., Carruthers, S. P., Berk, M., Rossell, S. L., & Van Rheenen, T. E. (2024). To what extent does white matter map to cognition in bipolar disorder? A systematic review of the evidence. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, **128**, 110868. https://doi.org/10.1016/j.pnpbp.2023.110868.

Cherbuin, N., Sachdev, P., & Anstey, K. J. (2010). Neuropsychological predictors of transition from healthy cognitive aging to mild cognitive impairment: The PATH through life study. *The American Journal of Geriatric Psychiatry*, **18**(8), 723–733. https://doi.org/10.1097/JGP.0b013e3181cdecf1.

Cho, H., Pilloni, G., Tahsin, R., Best, P., Krupp, L., Oh, C., & Charvet, L. (2023). Moving intra-individual variability (IIV) towards clinical utility: IIV measured using a commercial testing platform. *Journal of the Neurological Sciences*, **446**, 120586. https://doi.org/10.1016/j.jns.2023.120586.

Christensen, H., Dear, K. B., Anstey, K. J., Parslow, R. A., Sachdev, P., & Jorm, A. F. (2005). Within-occasion intraindividual variability and preclinical diagnostic status: Is intraindividual variability an indicator of mild cognitive impairment? *Neuropsychology*, **19**(3), 309.

Cornblatt, B. A., Risch, N. J., Faris, G., Friedman, D., & Erlenmeyer-Kimling, L. (1988). The continuous performance test, identical pairs version (CPT-IP): I. New findings about sustained attention in normal families. *Psychiatry Research*, **26**(2), 223–238. https://doi.org/10.1016/0165-1781(88)90076-5.

Cullen, B., Ward, J., Graham, N. A., Deary, I. J., Pell, J. P., Smith, D. J., & Evans, J. J. (2016). Prevalence and correlates of cognitive impairment in euthymic adults with bipolar disorder: A systematic review. *Journal of Affective Disorders*, **205**, 165–181. https://doi.org/10.1016/j.jad.2016.06.063.

Davis, J. J., Sivaramakrishnan, A., Rolin, S., & Subramanian, S. (2025). Intra-individual variability in cognitive performance predicts functional decline in Parkinson's disease. *Applied Neuropsychology: Adult*, **32**, 125–132. https://doi.org/10.1080/23279095.2022.2157276. Epub 2023 Jan 10.

Deary, I. J., & Der, G. (2005). Reaction time explains IQ'S association with death. *Psychological Science*, **16**(1), 64–69. https://doi.org/10.1111/j.0956-7976.2005.00781.x.

Depp, C. A., Savla, G. N., de Dios, L. A., Mausbach, B. T., & Palmer, B. W. (2012). Affective symptoms and intra-individual variability in the short-term course of cognitive functioning in bipolar disorder. *Psychological Medicine*, **42**(7), 1409–1416. https://doi.org/10.1017/s0033291711002662.

Dykiert, D., Der, G., Starr, J. M., & Deary, I. J. (2012). Age differences in intra-individual variability in simple and choice reaction time: Systematic review and meta-analysis. *PLoS One*, **7**(10), e45759. https://doi.org/10.1371/journal.pone.0045759.

Ferreira, D., Machado, A., Molina, Y., Nieto, A., Correia, R., Westman, E., & Barroso, J. (2017). Cognitive variability during middle-age: Possible association with neurodegeneration and cognitive reserve. *Frontiers in Aging Neuroscience*, **9**, 188. https://doi.org/10.3389/fnagi.2017.00188.

Fjell, A. M., Westlye, L. T., Amlien, I. K., & Walhovd, K. B. (2011). Reduced white matter integrity is related to cognitive instability. *The Journal of Neuroscience*, **31**(49), 18060–18072. https://doi.org/10.1523/jneurosci.4735-11.2011.

Fuermaier, A., Tucha, L., Koerts, J., Aschenbrenner, S., Kaunzinger, I., Hauser, J., & Tucha, O. (2015). Cognitive impairment in adult ADHD—Perspective matters! *Neuropsychology*, **29**(1), 45.

Gallagher, P., Nilsson, J., Finkelmeyer, A., Goshawk, M., Macritchie, K. A., Lloyd, A. J., & Watson, S. (2015). Neurocognitive intra-individual variability in mood disorders: Effects on attentional response time distributions. *Psychological Medicine*, **45**(14), 2985–2997. https://doi.org/10.1017/s0033291715000926.

Haatveit, B., Westlye, L. T., Vaskinn, A., Flaaten, C. B., Mohn, C., Bjella, T., & Ueland, T. (2023). Intra- and inter-individual cognitive variability in schizophrenia and bipolar spectrum disorder: An investigation across multiple cognitive domains. *Schizophrenia*, **9**(1), 89. https://doi.org/10.1038/s41537-023-00414-4.

Halliday, D. W., Gawryluk, J. R., Garcia-Barrera, M. A., & MacDonald, S. W. (2019). White matter integrity is associated with intraindividual variability in neuropsychological test performance in healthy older adults. *Frontiers in Human Neuroscience*, **13**, 352. https://doi.org/10.3389/fnhum.2019.00352.

Hultsch, D. F., MacDonald, S. W., & Dixon, R. A. (2002). Variability in reaction time performance of younger and older adults. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, **57**(2), P101–P115. https://doi.org/10.1093/geronb/57.2.p101.

Hultsch, D. F., & MacDonald, S. W. S. (2004). Intraindividual variability in performance as a theoretical window onto cognitive aging. In *New Frontiers in cognitive aging* (pp. 65–88). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198525691.003.0004.

Jackson, J. D., Balota, D. A., Duchek, J. M., & Head, D. (2012). White matter integrity and reaction time intraindividual variability in healthy aging and early-stage Alzheimer disease. *Neuropsychologia*, **50**(3), 357–366. https://doi.org/10.1016/j.neuropsychologia.2011.11.024.

Jahanshad, N., Kochunov, P. V., Sprooten, E., Mandl, R. C., Nichols, T. E., Almasy, L., & Glahn, D. C. (2013). Multi-site genetic analysis of diffusion images and voxelwise heritability analysis: A pilot project of the ENIGMA–DTI working group. *NeuroImage*, **81**, 455–469. https://doi.org/10.1016/j.neuroimage.2013.04.061.

Johnson, M. A., Diaz, M. T., & Madden, D. J. (2015). Global versus tract-specific components of cerebral white matter integrity: Relation to adult age and perceptual-motor speed. *Brain Structure and Function*, **220**(5), 2705–2720.

Jones, J. D., Burroughs, M., Apodaca, M., & Bunch, J. (2020). Greater intraindividual variability in neuropsychological performance predicts cognitive impairment in de novo Parkinson's disease. *Neuropsychology*, **34**(1), 24–30. https://doi.org/10.1037/neu0000577.

Kälin, A. M., Pflüger, M., Gietl, A. F., Riese, F., Jäncke, L., Nitsch, R. M., & Hock, C. (2014). Intraindividual variability across cognitive tasks as a potential marker for prodromal Alzheimer's disease. *Frontiers in Aging Neuroscience*, **6**, 147. https://doi.org/10.3389/fnagi.2014.00147.

Karantonis, J. A., Rossell, S. L., Carruthers, S. P., Sumner, P., Hughes, M., Green, M. J., & Van Rheenen, T. E. (2020). Cognitive validation of cross-diagnostic cognitive subgroups on the schizophrenia-bipolar spectrum. *Journal of Affective Disorders*, **266**, 710–721. https://doi.org/10.1016/j.jad.2020.01.123.

Kern, R. S., Nuechterlein, K. H., Green, M. F., Baade, L. E., Fenton, W. S., Gold, J. M., & Marder, S. R. (2008). The MATRICS consensus cognitive battery, part 2: Co-norming and standardization. *The American Journal of Psychiatry*, **165**(2), 214–220. https://doi.org/10.1176/appi.ajp.2007.07010043.

Krukow, P., Szaniawska, O., Harciarek, M., Plechawska-Wojcik, M., & Jonak, K. (2017). Cognitive inconsistency in bipolar patients is determined by increased intra-individual variability in initial phase of task performance. *Journal of Affective Disorders*, **210**, 222–225. https://doi.org/10.1016/j.jad.2016.12.050.

Li, X., Xu, M., & Wang, Z. (2023). Childhood trauma, intraindividual reaction time variability, baseline respiratory sinus arrhythmia, and perceived relapse tendency among males with substance use disorders. *The American Journal of Drug and Alcohol Abuse*, **49**(6), 827–838. https://doi.org/10.1080/00952990.2023.2289006.

Lövdén, M., Schmiedek, F., Kennedy, K. M., Rodrigue, K. M., Lindenberger, U., & Raz, N. (2013). Does variability in cognitive performance correlate with frontal brain volume? *NeuroImage*, **64**, 209–215. https://doi.org/10.1016/j.neuroimage.2012.09.039.

MacDonald, S. W., Hultsch, D. F., & Dixon, R. A. (2003). Performance variability is related to change in cognition: Evidence from the Victoria longitudinal study. *Psychology and Aging*, **18**(3), 510–523. https://doi.org/10.1037/0882-7974.18.3.510.

MacDonald, S. W., Li, S. C., & Bäckman, L. (2009). Neural underpinnings of within-person variability in cognitive functioning. *Psychology and Aging*, **24**(4), 792–808. https://doi.org/10.1037/a0017798.

MacDonald, S. W. S., Nyberg, L., & Bäckman, L. (2006). Intra-individual variability in behavior: Links to brain structure, neurotransmission and neuronal activity. *Trends in Neurosciences*, **29**(8), 474–480. https://doi.org/10.1016/j.tins.2006.06.011.

Madden, D. J., Bennett, I. J., & Song, A. W. (2009). Cerebral white matter integrity and cognitive aging: Contributions from diffusion tensor imaging. *Neuropsychology Review*, **19**(4), 415–435. https://doi.org/10.1007/s11065-009-9113-2.

Mazerolle, E. L., Wojtowicz, M. A., Omisade, A., & Fisk, J. D. (2013). Intra-individual variability in information processing speed reflects white matter microstructure in multiple sclerosis. *Neuroimage Clinical*, **2**, 894–902. https://doi.org/10.1016/j.nicl.2013.06.012.

Mella, N., de Ribaupierre, S., Eagleson, R., & de Ribaupierre, A. (2013). Cognitive Intraindividual variability and white matter integrity in aging. *The Scientific World Journal*, **2013**, 350623. https://doi.org/10.1155/2013/350623.

Mitchell, A. J., Kemp, S., Benito-León, J., & Reuber, M. (2010). The influence of cognitive impairment on health-related quality of life in neurological disease. *Acta Neuropsychiatrica*, **22**(1), 2–13. https://doi.org/10.1111/j.1601-5215.2009.00439.x.

Montgomery, S. A., & Åsberg, M. (1979). A new depression scale designed to be sensitive to change. *British Journal of Psychiatry*, **134**(4), 382–389. https://doi.org/10.1192/bjp.134.4.382.

Mori, S., Oishi, K., Jiang, H., Jiang, L., Li, X., Akhter, K., & Woods, R. (2008). Stereotaxic white matter atlas based on diffusion tensor imaging in an ICBM template. *NeuroImage*, **40**(2), 570–582.

Moss, R. A., Finkelmeyer, A., Robinson, L. J., Thompson, J. M., Watson, S., Ferrier, I. N., & Gallagher, P. (2016). The impact of target frequency on intra-individual variability in euthymic bipolar disorder: A comparison of two sustained attention tasks. *Frontiers in Psychiatry*, **7**, 106. https://doi.org/10.3389/fpsyt.2016.00106.

Moy, G., Millet, P., Haller, S., Baudois, S., de Bilbao, F., Weber, K., … Delaloye, C. (2011). Magnetic resonance imaging determinants of intraindividual variability in the elderly: Combined analysis of grey and white matter. *Neuroscience*, **186**, 88–93. doi: https://doi.org/10.1016/j.neuroscience.2011.04.028

Neill, E., & Rossell, S. L. (2013). Executive functioning in schizophrenia: The result of impairments in lower order cognitive skills? *Schizophrenia Research*, **150**(1), 76–80. https://doi.org/10.1016/j.schres.2013.07.034.

Nilsson, J., Thomas, A. J., O'Brien, J. T., & Gallagher, P. (2014). White matter and cognitive decline in aging: A focus on processing speed and variability. *Journal of the International Neuropsychological Society*, **20**(3), 262–267. https://doi.org/10.1017/S1355617713001458.

Nuechterlein, K. H., Green, M. F., Kern, R. S., Baade, L. E., Barch, D. M., Cohen, J. D., & Gold, J. M. (2008). The MATRICS consensus cognitive battery, part 1: Test selection, reliability, and validity. *American Journal of Psychiatry*, **165**(2), 203–213.

Rajji, T. K., Miranda, D., & Mulsant, B. H. (2014). Cognition, function, and disability in patients with schizophrenia: A review of longitudinal studies. *The Canadian Journal of Psychiatry*, **59**(1), 13–17. https://doi.org/10.1177/070674371405900104.

Ram, N., Rabbitt, P., Stollery, B., & Nesselroade, J. R. (2005). Cognitive performance inconsistency: Intraindividual change and variability. *Psychology and Aging*, **20**, 623–633. https://doi.org/10.1037/0882-7974.20.4.623.

Reitan, R. M. (1958). Validity of the trail making test as an indicator of organic brain damage. *Perceptual and Motor Skills*, **8**(3), 271–276. https://doi.org/10.2466/pms.1958.8.3.271.

Roalf, D. R., Rupert, P., Mechanic-Hamilton, D., Brennan, L., Duda, J. E., Weintraub, D., & Moberg, P. J. (2018). Quantitative assessment of finger tapping characteristics in mild cognitive impairment Alzheimer's disease, and Parkinson's disease. *Journal of Neurology*, **265**(6), 1365–1375. https://doi.org/10.1007/s00415-018-8841-8.

Sánchez-Torres, A. M., García de Jalón, E., Gil-Berrozpe, G. J., Peralta, V., & Cuesta, M. J. (2023). Cognitive intraindividual variability, cognitive impairment and psychosocial functioning in first-episode psychosis patients. *Psychiatry Research*, **328**, 115473. https://doi.org/10.1016/j.psychres.2023.115473.

Sheehan, D.V., Lecrubier, Y., Sheehan, K.H., Amorim, P., Janavs, J., Weiller, E., … Dunbar, G.C. (1998). The Mini-international neuropsychiatric interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *The Journal of Clinical Psychiatry*, **59** *Suppl* 20, 22–33;quiz 34–57.

Silbert, L. C., Dodge, H. H., Perkins, L. G., Sherbakov, L., Lahna, D., Erten-Lyons, D., & Kaye, J. A. (2012). Trajectory of white matter hyperintensity burden preceding mild cognitive impairment. *Neurology*, **79**(8), 741–747. https://doi.org/10.1212/WNL.0b013e3182661f2b.

Simonsen, C., Sundet, K., Vaskinn, A., Ueland, T., Romm, K. L., Hellvin, T. Melle, I., Friis, S., & Andreassen, O. A. (2010). Psychosocial function in schizophrenia and bipolar disorder: Relationship to neurocognition and clinical symptoms. *Journal of the International Neuropsychological Society*, **16** (5), 771–783. https://doi.org/10.1017/s1355617710000573.

Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E.J., Johansen-Berg, H., … Matthews, P.M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, **23** (Suppl. 1), S208–S219. https://doi.org/10.1016/j.neuroimage.2004.07.051

Tamnes, C. K., Fjell, A. M., Westlye, L. T., Østby, Y., & Walhovd, K. B. (2012). Becoming consistent: Developmental reductions in intraindividual variability in reaction time are related to white matter integrity. *Journal of Neuroscience*, **32**(3), 972–982.

Tan, E. J., & Rossell, S. L. (2014). Building a neurocognitive profile of thought disorder in schizophrenia using a standardized test battery. *Schizophrenia Research*, **152**(1), 242–245. https://doi.org/10.1016/j.schres.2013.11.001.

Tucker-Drob, E. M. (2009). Differentiation of cognitive abilities across the life span. *Developmental Psychology*, **45**(4), 1097–1118. https://doi.org/10.1037/a0015864.

Van Rheenen, T. E., Lewandowski, K. E., Bauer, I. E., Kapczinski, F., Miskowiak, K., Burdick, K. E., & Balanzá-Martínez, V. (2020). Current understandings of the trajectory and emerging correlates of cognitive impairment in bipolar disorder: An overview of evidence. *Bipolar Disorders*, **22**(1), 13–27. https://doi.org/10.1111/bdi.12821.

Van Rheenen, T. E., & Rossell, S. L. (2014a). An empirical evaluation of the MATRICS consensus cognitive battery in bipolar disorder. *Bipolar Disorders*, **16**(3), 318–325. https://doi.org/10.1111/bdi.12134.

Van Rheenen, T. E., & Rossell, S. L. (2014b). Investigation of the component processes involved in verbal declarative memory function in bipolar disorder: Utility of the Hopkins verbal learning test-revised. *Journal of the International Neuropsychological Society*, **20**(7), 727–735. https://doi.org/10.1017/s1355617714000484.

Wechsler, D. (2001). *Wechsler test of adult reading: WTAR*. Psychological Corporation.

Williams, B. R., Hultsch, D. F., Strauss, E. H., Hunter, M. A., & Tannock, R. (2005). Inconsistency in reaction time across the life span. *Neuropsychology*, **19**(1), 88–96. https://doi.org/10.1037/0894-4105.19.1.88.

Wojtowicz, M., Omisade, A., & Fisk, J. D. (2013). Indices of cognitive dysfunction in relapsing-remitting multiple sclerosis: Intra-individual variability, processing speed, and attention network efficiency. *Journal of the International Neuropsychological Society*, **19**(5), 551–558. https://doi.org/10.1017/s1355617713000027.

Yatham, L. N., Torres, I. J., Malhi, G. S., Frangou, S., Glahn, D. C., Bearden, C. E., & Chengappa, K. N. R. (2010). The International Society for Bipolar Disorders–Battery for assessment of Neurocognition (ISBD-BANC). *Bipolar Disorders*, **12**(4), 351–363. https://doi.org/10.1111/j.1399-5618.2010.00830.x.

Young, R. C., Biggs, J. T., Ziegler, V. E., & Meyer, D. A. (1978). A rating scale for mania: Reliability, validity and sensitivity. *British Journal of Psychiatry*, **133**(5), 429–435. https://doi.org/10.1192/bjp.133.5.429.

https://doi.org/10.1017/S0033291725100731 Published online by Cambridge University Press