Title: Slow binocular rivalry: Translating a replicated biomarker for bipolar disorder

Running title: Translating binocular rivalry rate

Phillip C. F. Law[1], Steven M. Miller[1*], Trung T. Ngo[2*]

[1] Monash Biomedicine Discovery Institute, Department of Physiology, Monash University, Melbourne, Australia.

[2] UQ Diamantina Institute, University of Queensland Faculty of Medicine & Translational Research Institute, Woolloongabba, Brisbane, QLD, Australia

[*] Steven M. Miller and Trung T. Ngo are co-senior authors.

Corresponding author:

Phillip C. F. Law

phillip.law@monash.edu

Abstract

Biomarkers for psychiatric disorders have not yet achieved clinical translation. Two reasons for this are use of small study samples and a lack of standardized test protocols. One psychiatric biomarker — slow binocular rivalry (BR) in bipolar disorder (BD) — has been replicated in small studies but not translated. Psychiatric heterogeneity suggests clinical translation studies require large samples (N=100s–1000s) and genetic translation studies even larger, indeed massive samples (N=10,000s–100,000s+). Translational biomarker studies therefore present enormous logistical and economic challenges. We therefore proposed, and here investigate, an efficient low-cost online BR test platform for large-scale testing of existing well-characterized psychiatric participant cohorts. The platform utilizes inexpensive cardboard-framed anaglyph filter (AF) glasses that can be readily mailed to subjects for at-home testing. It additionally enables protocol standardization for all translational BR studies, whether testing occurs at home or in the laboratory. In a laboratory-based study, we compared BR with red/blue AFs and a previously-used monochromatic (green) presentation method (polarization filters; PFs) in 20 outpatient persons with BD and 20 age- and sex-matched healthy controls. We also varied stimulus drift speed to examine stimulus optimization. Red/blue-AF gratings drifting at 4 cycles/s were optimal according to key criteria and elicited slow reliable BR in persons with BD, comparable to the green-PF gratings used in previous BR clinical and heritability studies. This study provides psychophysical validation for a test method that standardizes laboratory-based BR testing and that can also be used online to logistically achieve the large samples required for clinical and genetic translational studies.

Keywords: binocular rivalry rate endophenotype, genome-wide association studies, bipolar disorder, anaglyphs, online test method

Introduction

Despite the rise in biological psychiatry research in recent decades, there are no routinely used psychiatric diagnostic tests available to clinicians. Beyond traditional clinical interview, there are no means to improve diagnosis of psychiatric disorders according to classical categories, to identify disorder subtypes within classical categories, to identify biologically similar subtypes beyond classical categories (e.g., aligning with Research Domain Criteria; RDoC), or to predict risk, prognosis or response to particular treatments[1] (notwithstanding routine plasma lithium level testing to ensure optimal and safe lithium administration[2]). Reasons for this state of affairs have been examined via white papers[3,4] and think tanks[1] and two important such reasons are (i) use of small study sample sizes, and (ii) a lack of research protocol standardization.

Evidence has been gathered to support the proposal that slow binocular rivalry (BR) is a biomarker for the heritable psychiatric condition, bipolar disorder[5-7] (BD). BR is the involuntary perceptual alternation of conflicting images presented simultaneously to each eye in the same retinal location. BR has been studied for more than 100 years including psychophysical, electrophysiological, neuroimaging, neurostimulation, computational, clinical and pharmacological approaches[8-12]. Although there were reports of BR rate (BRR) anomalies in clinical psychiatric groups in the mid-20th Century, the modern clinical focus on BR was stimulated by reports that BRR is slow in persons with BD compared with controls[5,7] (CTL). This finding in an Australian cohort of 50 BD-I subjects and 93 CTL subjects has since been replicated in cohorts in Japan[13] (N=11 BD-I, 17 BD-II, 25 CTL), New Zealand[14] (N=71 BD-I, 24 CTL) and China[15] (N=14 BD-I, 30 CTL).

The slow BR trait is reported to be: (i) highly reliable[5,7,16]; (ii) relatively unaffected by BD medications[5,7,13,14] (though see Jia et al.[17] and Discussion); (iii) unaffected by cognitive functioning in persons with BD[14]; (iv) evident even when persons with BD are clinically

well[6,14] (though see Discussion); and (v) not explained by eye movements[18,19]. Moreover, a large twin study reported that individual variation in BRR is ~50% genetically determined[16] (N=722; see also Shannon et al.[20]). Together, these findings support the proposal that slow BR may be a useful biomarker (or endophenotype) to increase power in BD genetic studies[6,7,16,21]. While large-scale consortium genome-wide association studies (GWAS) have identified genes predisposing to BD[22,23], the complex genetic basis of BD, as with all psychiatric disorders, remains poorly understood. Moreover, GWAS require massive sample sizes, in the order of tens of thousands, even hundreds of thousands or more[24]. Consequently, there are very substantial logistics and cost burdens of using laboratory-based BR testing as an endophenotype for BD GWAS (a problem even more significant for laboratory-based endophenotype studies using techniques such as neuroimaging, eye movement recording, or electrophysiology; see Ivleva et al.[25]). To overcome such challenges, we proposed the notion of an *online at-home* BR test platform[26], as detailed further below.

Finding a statistically significant difference between clinical psychiatric and CTL groups does not in itself present opportunity for clinical translation of a biomarker. Potential diagnostic tests need to demonstrate utility in distinguishing diagnostically unclear real-world presentations[1,4]. In the case of persons with BD, misdiagnosis is common[27-31] and a diagnostic test to aid discrimination would be helpful, such as distinguishing psychosis in persons with BD from that of persons with schizophrenia (SCZ) or the depressive phase of BD from that of major depressive disorder[6] (MDD). Accurate diagnosis in these contexts would direct treatment choice and reduce treatment change, leading to more focused treatment options. However, studies of BRR in SCZ have reported conflicting findings. One reported no BRR difference between SCZ and CTL groups[5] (N=18 SCZ, 30 CTL) while another reported significantly slower BR in the SCZ group[32] (N=40 SCZ, 40 CTL). Similarly, conflicting results have been reported for BRR in persons with MDD, including no difference from CTLs[5] (N=18 MDD, 30

CTL) and slower BRR in MDD (N=28 MDD, 30 CTL in Jia et al.[17]; N=16 MDD, 30 CTL in Zhu et al.[15]). Larger sample sizes, in the order of hundreds to thousands of subjects, are needed to clarify such discrepancies and to address clinical translation prospects of the BR biomarker. Moreover, no studies have examined diagnostic discrimination prospectively or whether BRR can be used to biologically stratify psychiatric patients within or beyond classical diagnostic boundaries. Additionally, no studies have examined whether BRR can predict prognosis, risk of developing psychiatric disorder in vulnerable populations, or particular treatment responsiveness.

While GWAS of BD using BRR as an endophenotype are yet to be performed, there have nonetheless been two BR GWAS (N=1,051 in Bosten et al.[33] and N=2,097 in Chen et al.[34]), but with conflicting genetic association findings. Though these studies had impressive sample sizes by psychophysical standards, they were small by GWAS standards and the authors of both acknowledged the need for larger samples. Moreover, these GWAS differed substantially from each other in the BR test protocols they used and both differed substantially from the protocols used in the original BR clinical and heritability studies[5,7,16]. Differing BR test protocols (along with small sample sizes) may have also accounted for the conflicting reports of BRR in SCZ and MDD mentioned earlier. Common protocol differences between studies include total BR viewing time (and hence whether analyzed BRRs are stabilized), whether mixed percepts are recorded and excluded from BRR calculation, and different BR stimulus types. Because all such factors can influence BRR, clinical and genetic BR studies would benefit from a test method that provides a standardized test protocol, in addition to providing the means for feasibly obtaining large and even massive sample sizes.

Such a BR presentation method involves mail-able cardboard-framed glasses with anaglyph filters (AFs) and performing BR online[26]. This method enables inexpensive, cost-effective, accessible BR testing with a standardized test protocol. In the case of at-home testing

this can occur with well-characterized patient and CTL groups drawn from existing large-scale clinical and psychiatric genetic cohorts via international multicentre collaboration. At-home BR testing can feasibly obtain the massive sample sizes required for GWAS and large sample sizes required for clinical discrimination studies. This is because existing research programs with such cohorts typically have restrictions on in-laboratory assessment time per participant and collaborations that add to such time are less likely to be established. Conversely, collaborations that do not encumber in-laboratory assessment time (such as use of at-home phenotyping) are more logistically feasible and thus more likely to be established. Importantly, such collaborations are able to capitalize on the substantial resources already invested in clinical and genetic characterization of existing cohorts.

However, psychophysical validation data are required to support this proposed online BR test method. Although red/blue AFs elicit excellent BR with the least bleeding of each eye's image into the other[26], it is not clear whether an individual's BRR with different colored (red/blue) orthogonal gratings in each eye is comparable to that with our previously-used presentation methods (polarization filters, PFs; liquid crystal shutters, LCS) with orthogonal gratings of the same color (green) in each eye. There are reports of color effects on BRR — one reported significantly fewer alternations (slower BR) for green orthogonal gratings compared with red and red-green orthogonal gratings[35], while another reported no difference between red-green and black-white orthogonal gratings[36,37] — suggesting stimulus color differences between the AFs and PFs/LCS methods could potentially confound BRR data.

In addition, stimulus characteristics require optimization for genetic and clinical studies. In BR literature, the *strength* of a stimulus is the degree to which the physical characteristics of one eye's stimulus induce perceptual suppression of the other eye's stimulus[38]. Comparing data from two previous studies of BRR in BD that used different stimulus strengths in separate sets of subjects — i.e., one using low-strength stimuli[5] (green

stationary gratings with spatial frequency of 4 cycles/°) and the other using high-strength stimuli[7] (green gratings drifting at 4 cycles/s with spatial frequency of 8 cycles/°) — suggested the high-strength stimuli may better distinguish persons with BD from CTLs by driving CTLs' BRR faster (discussed in Miller et al.[5]). We recently examined stimulus-related BRR modulation in CTLs using PFs[39] and found that our previously-used BR grating stimuli drifting at 4 cycles/s may be less optimal than gratings drifting faster at 8 cycles/s. However, BRRs elicited with these stimulus types were only minimally different and we therefore concluded that the issue of stimulus optimization required direct assessment in a BD cohort.

On this background, we here report BRR elicited with red/blue-AFs and green-PFs in age- and sex-matched BD and CTL groups, and with varying grating drift speeds. We aimed to provide psychophysical validation data for the use of red/blue-AFs — which are suitable for large-scale online BR studies — and to determine optimal grating drift speed for clinical and genetic translational BR studies.

Methods

Twenty experiment-naïve clinically stable outpatient persons with BD (12 males; mean age=41.6 ± 13.4 years) and 20 age- and sex-matched healthy CTLs (mean age=41.6 ± 14.1 years) were recruited from the Monash Alfred Psychiatry Research Centre participant database. All participants including CTLs had their diagnostic status confirmed according to Diagnostic and Statistical Manual of Mental Disorders (4^{th} edition) with the Mini International Neuropsychiatric Interview[40]. Persons with BD were screened to exclude those with co-morbid psychiatric conditions, and CTLs were screened to exclude individuals with a psychiatric disorder (e.g., BD, SCZ, MDD) or first-degree relatives diagnosed with a psychiatric disorder. Participant assessments and screening, and the running of all experiments, was performed by author PCFL who was not blind to diagnosis at the time of BR testing. All participants were

reimbursed AU$40, AU$30, and AU$30 for sessions one, two, and three, respectively.

A full description of participant exclusion criteria, demographics, clinical features, medication status, premorbid intelligence and clinical state-trait instruments and scores (Montgomery-Åsberg Depression Rating Scale, MÅDRS; Young Mania Rating Scale, YMRS; subjective mood 10-point self-report visual analog scale, VAS; State-Trait Anxiety Inventory, STAI; Wechsler Test of Adult Reading, WTAR) has been published in detail previously in a study regarding eye movements and BRR[18] that used the same cohort and testing sessions as the current study.

Similarly, a full description of general experimental procedures, experimental equipment, stimulus specifications, BR testing procedures and response protocols, and BR data processing procedures was published previously[18]. The full within-subjects experimental protocol is visualized in Figure 1. Briefly, the stimuli were (i) stationary green-PF orthogonal gratings, (ii) drifting green-PF orthogonal gratings of 4 and 8 cycles/s, and (iii) drifting red/blue-AF orthogonal gratings or rings of 4 and 8 cycles/s. Only drifting stimuli were retested on a separate day. Each stimulus was viewed in a 7-min block comprising four 100-s trials (the blocks were separated by 110-s breaks and the trials by 30-s breaks). Participants were instructed to press one of three adjacent keyboard buttons in response to the left eye's image, the right eye's image, and the third response option being for either mixed or unusual percepts or to indicate a previous response error. BRR was calculated as the total number of perceptual alternations divided by the total time of BR viewing (expressed in Hz), excluding the mixed-percept responses.

The mixed-percept duration (MPD) measure is an approximation of the total time spent perceiving mixed percepts due to its conflation with previous erroneous response. With some degree of noise due to this conflation, a shorter MPD may nonetheless indicate a relatively greater amount of data on which BRR calculation is made, thus reflecting a more accurate and

representative measure of an individual's true BRR. This measure is certainly more likely to provide an individual's accurate BRR than test protocols in which the mixed percepts are, by instruction, subsumed into one or the other reported exclusive percept (an instruction that would artificially lengthen exclusive percept dominance durations and thus artificially slow BRR).

Written, informed consent was obtained from all participants in the presence of a witness. The study conformed to the Declaration of Helsinki and was approved by Alfred Human Research Ethics Committee and Monash University Human Research Ethics Committee.

For the data presented in the current study, normality was violated for the distributions of BRR, MPD, STAI-state, MÅDRS, YMRS, VAS, and WTAR (Shapiro-Wilk, $p \leq 0.05$), therefore non-parametric statistics were performed for these measures (e.g., Friedman test, Mann-Whitney U test, Wilcoxon signed rank test, Spearman's ρ, Kruskal-Wallis H-test). Planned and exploratory correlational analyses were one-tailed and two-tailed, respectively.

[Insert Figure 1 about here; Figure 1 in color]

Results

Slow BR in persons with BD

Table 1 shows that BRR was significantly slower in the BD group than in the CTL group for 4 and 8 cycles/s green-PF gratings, 4 and 8 cycles/s red/blue-AF rings, and 4 cycles/s red/blue-AF gratings ($p \leq 1.44 \times 10^{-3}$; $\alpha_{\text{Bonferroni-adjusted}}$: 0.05/7 Mann-Whitney U tests$=7.14 \times 10^{-3}$). However, there was no significant group difference in BRR for 8 cycles/s red/blue-AF gratings or for stationary green-PF gratings ($p \geq 1.22 \times 10^{-2}$). These results indicate slower BRR in persons with BD compared with CTLs for most of the stimulus conditions, with just one drifting

stimulus and the stationary gratings not showing this finding. The results confirm that: (i) the AF BR presentation method — which can be used in online BR studies — elicits slow BR in BD with gratings drifting at 4 cycles/s or rings at either speed; and (ii) in accordance with predictions, BR with the lower strength (stationary) gratings does not adequately distinguish persons with BD from CTL subjects.

[Insert Table 1 about here]

Comparison and correlation of BRRs between green-PF and red/blue-AF stimuli

Pairwise (within-group) comparisons showed, in both BD and CTL groups, no significant difference in BRR between same-speed green-PF gratings and red/blue-AF stimuli ($p \geq 1.69\times10^{-2}$; $\alpha_{\text{Bonferroni-adjusted}}$: 0.05/4 Wilcoxon signed rank tests=1.25×10^{-2}). Planned correlations of BRRs between same-speed green-PF and red/blue-AF stimulus conditions in persons with BD and CTLs were positive in both subject groups for all stimulus conditions. In both BD and CTL groups, there were moderate-to-high correlations between: (i) 4 cycles/s green-PF gratings and red/blue-AF gratings (BD: ρ=0.61, CTL: ρ=0.87; $p \leq 2.25\times10^{-3}$; $\alpha_{\text{Bonferroni-adjusted}}$: 0.05/4 Spearman's ρ tests=1.25×10^{-2}); and (ii) 4 cycles/s green-PF gratings and red/blue-AF rings (BD: ρ=0.56, CTL: ρ=0.79; $p \leq 4.92\times10^{-3}$). In the CTL group, there was also a significant high correlation between 8 cycles/s green-PF gratings and red/blue-AF gratings (ρ=0.91; $p=8.66\times10^{-9}$), and a significant moderate correlation between 8 cycles/s green-PF gratings and red/blue-AF rings (ρ=0.56; $p=4.80\times10^{-3}$). However, such correlations were not significant in the BD group ($p \geq 2.46\times10^{-2}$). Overall these results indicate BRR elicited with red/blue-AF gratings is comparable to that with same-speed green-PF gratings, in both the BD and CTL groups. Moreover, the results indicate positive moderate-to-high association in BD individuals' and CTL individuals' BRRs for most stimuli, including between our previously-

used 4 cycles/s green-PF gratings and same-speed red/blue-AF gratings that are suitable for use in online BR studies.

Retest reliability of BRR

Planned correlations between test-retest BRRs were positive for all stimulus conditions in both subject groups. Table 1 shows a significant, moderate-to-high correlation in BRR between testing sessions for (i) 4 cycles/s green-PF gratings and (ii) 4 and 8 cycles/s red/blue-AF gratings in both subject groups (BD: ρ=0.73–0.79, CTL: ρ=0.71–0.83; $p \leq 2.02\times10^{-4}$; $\alpha_{\text{Bonferroni-adjusted}}$: 0.05/6 Spearman's ρ tests=8.33×10^{-3}). In addition, there was a significant, high correlation in BRR between testing sessions for (i) 8 cycles/s green-PF gratings and (ii) 4 and 8 cycles/s red/blue-AF rings in the CTL group (ρ=0.75–0.83, respectively; $p \leq 6.69\times10^{-5}$), but not in the BD group ($p \geq 1.71\times10^{-2}$). These results indicate moderate-to-high reliability of BRR between different days in both persons with BD and CTLs for 4 cycles/s green-PF gratings, consistent with our previous studies for CTLs[16] and boosting our early BD retest data with these stimuli[7]. Moreover, the results indicate moderate-to-high retest reliability of BRR in persons with BD and CTLs for the 4 cycles/s red/blue-AF gratings, but not for AF rings at either speed.

Stimulus strength modulation of BRR

We compared BRRs between drift speeds for each stimulus type in persons with BD and CTLs. A Friedman test with BRR as the dependent variable and stimulus condition as the independent variable showed a significant difference in BRR across the stimulus conditions in both BD and CTL groups ($p \leq 9.74\times10^{-4}$). Pairwise (within-group) comparisons showed a signficantly faster BRR for 4 and 8 cycles/s green-PF gratings compared with stationary green-PF gratings in the CTL group ($p \leq 3.90\times10^{-4}$; $\alpha_{\text{Bonferroni-adjusted}}$: 0.05/5 Wilcoxon signed rank tests=0.01), but not in

the BD group ($p \geq 0.60$). BRRs between 4 and 8 cycles/s for all stimulus types were not significantly different in both the BD and CTL groups ($p \geq 0.25$). These results indicate that maximizing BD and CTL group separation benefits from the rivaling gratings having a motion component and that this component can be with drift speeds of either 4 or 8 cycles/s.

Furthermore, as observed in Table 1, the greater group separation with moving gratings is driven by a faster BRR with these stimuli in the CTL group, while BRR in the BD group appears to remain robustly slow irrespective of stimulus type, i.e., persons with BD appear to exhibit a relative insensitivity to stimulus-strength modulation, in accordance with the prediction by Miller et al.[5]. However, the results regarding stimulus strength modulation of BRR in both BD and CTL groups are in fact not as clear as the central tendency values suggest when examined from an individual variation perspective. These data are presented in detail in a separate report that addresses individual variation findings for stimulus strength modulation in the context of classical psychophysics principles of BR dynamics (Law et al., in preparation).

MPD

Table 2 shows that MPD was significantly shorter in the CTL group than in the BD group for 8 cycles/s green-PF gratings ($p = 4.68 \times 10^{-3}$; $\alpha_{\text{Bonferroni-adjusted}}$: 0.05/7 Mann-Whitney U tests$=7.14 \times 10^{-3}$), but not for the remaining stimulus conditions ($p \geq 3.04 \times 10^{-2}$). Thus we compared MPDs between stimulus conditions in persons with BD and CTLs. A Friedman test with MPD as the dependent variable and stimulus condition as the independent variable showed a significant difference in MPD across the stimulus conditions in both BD and CTL groups ($p \leq 1.90 \times 10^{-5}$). Pairwise (within-group) comparisons showed, in both BD and CTL groups, a significantly shorter MPD for 4 and 8 cycles/s red/blue-AF gratings than for 8 cycles/s green-PF gratings ($p \leq 4.49 \times 10^{-4}$; $\alpha_{\text{Bonferroni-adjusted}}$: 0.05/21 Wilcoxon signed rank tests$=2.38 \times 10^{-3}$). MPD was also significantly shorter for 4 cycles/s red/blue-AF gratings than for stationary and

4 cycles/s green-PF gratings in the CTL group ($p \leq 1.51 \times 10^{-3}$), but not in the BD group ($p \geq 3.31 \times 10^{-3}$). However, MPD was not significantly different between 4 and 8 cycles/s red/blue-AF gratings in either the BD or CTL group ($p \geq 7.31 \times 10^{-2}$). Overall these results indicate that in both persons with BD and CTLs, red/blue-AF gratings drifting at 4 cycles/s — which elicit reliably slow BRR in BD — yielded a short MPD and thus a more accurate measure of an individual's true BRR (notwithstanding interpretation issues; see Discussion).

[Insert Table 2 about here]

Retest reliability of MPD

Exploratory correlations between test-retest MPDs were positive for all stimulus conditions in both subject groups. Table 2 shows a significant, moderate-to-high correlation in MPD between testing sessions for all drifting stimuli in the CTL group ($\rho = 0.58–0.81$; $p \leq 7.55 \times 10^{-3}$; $\alpha_{\text{Bonferroni-adjusted}}$: 0.05/6 Spearman's ρ tests=8.33×10^{-3}). In the BD group, there was a significant, moderate-to-high correlation in MPD between testing sessions for most drifting stimuli ($\rho = 0.62–0.72$; $p \leq 3.39 \times 10^{-3}$) with the exception of 4 cycles/s red/blue-AF gratings and rings ($p \leq 2.88 \times 10^{-2}$). These results indicate moderate-to-high reliability of MPD between different days in both persons with BD and CTLs for 4 and 8 cycles/s green-PF gratings and 8 cycles/s red/blue-AF stimuli, but not for 4 cycles/s red/blue-AF stimuli (notwithstanding interpretation issues; see Discussion).

Association between BR measures, clinical state, psychometric measures, age and medication

We conducted planned and exploratory correlational analyses between the BR measures (BRR, MPD) and subjects' age along with each of their psychometric measures (STAI-trait, STAI-

state, MÅDRS, YMRS, VAS, WTAR) for all stimuli tested. In both BD and CTL groups, age and psychometric measures were not significantly correlated with BRR ($p \geq 2.31 \times 10^{-2}$, one-tailed; $\alpha_{\text{Bonferroni-adjusted}}$: 0.05/35 Spearman's ρ tests=1.43×10^{-3}) or MPD ($p \geq 3.78 \times 10^{-3}$, two-tailed). There was also no significant effect of medication on BRR or MPD in the BD group (Kruskal Wallis test; $p \geq 0.15$). Overall these results indicate that the BR measures examined are not influenced by age, subject factors, and medication.

Discussion

The current study examined whether BRR induced by red/blue-AF stimuli — which are suitable for large-scale online BR endophenotype studies involving tens to hundreds of thousands of subjects — is comparable to that with previously-used green-PF gratings. It also compared gratings with different drift speeds to examine which are optimal for large-scale studies, on a variety of key criteria (Table 3). We again replicated the finding of slow BRR in persons with BD relative to CTLs[5,7,13-15] for all but one of the drifting stimuli and the stationary gratings. Importantly for future large-scale clinical and genetic translational BR studies, we found that: (i) BRR elicited with red/blue-AF gratings is comparable to that with previously-used green-PF gratings, at the same drift speed of 4 cycles/s, in both the BD and CTL groups; and accordingly, (ii) red/blue-AF stimuli at this drift speed induces reliable BRR that is significantly slower in persons with BD than CTLs.

The mean BRRs elicited with 4 cycles/s red/blue-AF gratings in persons with BD and CTLs (0.35Hz and 0.54Hz, respectively) were not significantly different to those of same-speed green-PF gratings in the current study (0.30Hz and 0.51Hz, respectively), and were comparable to BRRs reported in our previous studies with green gratings viewed through either PFs or LCS glasses[5,16,39] (0.31Hz and 0.53–0.60 Hz, respectively). A positive moderate-to-high correlation between BRRs elicited with 4 cycles/s red/blue-AF and green-PF gratings was also observed

in BD and CTL groups. Moreover, the moderate-to-high between-session reliability of BRR reported previously for green gratings viewed through LCS[7,16] was also observed in the current study with 4 cycles/s red/blue-AF gratings in both BD and CTL groups. Together, these findings show that the color differences examined do not affect BRR in either persons with BD or CTLs. This provides foundational psychophysical validation data for the online BR test platform. Furthermore, the findings comparing stimulus drift speeds provide stimulus optimization data for large-scale online BR studies and standardized laboratory-based BR studies and suggest (Table 3) that optimal stimuli are red/blue-AF gratings drifting at 4 cycles/s (i.e., the same drift speed as used in our previous studies).

[Insert Table 3 about here]

The findings of MPD as a secondary measure of interest also support the use of 4 cycles/s red/blue-AF gratings for online BR studies. However, MPD data should be interpreted with caution due to the conflation of MPD responses with reporting of erroneous responses in our test protocol. Thus, rather than MPD per se varying according to presentation method, it may be that the erroneous response rate varies in this way (i.e., fewer errors for 4 cycles/s red/blue-AF gratings in persons with BD and CTLs could also explain apparently shorter MPD for that stimulus condition). Even if so, it would still be true that this 4 cycles/s stimulus yielded a more accurate measure of an individual's true BRR (i.e., be it through fewer errors, a shorter actual MPD, or a combination of these factors). In our experience however, error responses tend to be infrequent across the long total viewing times and are thus unlikely to meaningfully affect obtained BRRs. Nonetheless, the online BR test platform includes a non-rivalrous catch trial component to identify and exclude subjects with high error rates (detailed further below).

In both BD and CTL groups, BRR and MPD were not influenced by age, depressive or

manic symptom severity, state anxiety, trait anxiety, subjective mood, premorbid intelligence, or medication. These findings are consistent with: (i) our previous study of healthy subjects[39]; (ii) a study showing BRR in healthy subjects is not influenced by state anxiety[41]; and (iii) reports that BRR in persons with BD is not influenced by depressive or manic symptoms[5,14] or medication[5,13,14] (though see Carter et al.[42]; Nagamine et al.[43]; van Loon et al.[44]). However, the present findings conflict with reports that BRR is associated with age[33,45], trait anxiety[41,46], and depressive or manic symptom severity[15,17]. These discrepancies in subject factor findings may have arisen due to small sample sizes and differing test protocols, underscoring the need for large-scale and standardized BR testing. Ultimately, rigorous examination of state and medication effects will require BRR testing before and after state or medication change/onset while controlling for their mutual confounding effects[5,6]. Further relevance of the stimulus strength modulation findings observed in the current study are also the subject of a separate report addressing classical psychophysics principles of BR dynamics (Law et al., in preparation).

There are further issues that will need to be examined before the online test platform is fully validated and these are being addressed as development of the platform progresses (see www.binocularrivalryonline.com). Importantly though, the platform is premised on the notion of *minimal phenotyping* (discussed in Law et al.[26]). According to this notion, small inaccuracies (noise) introduced by using cost-effective phenotyping methods (such as at-home testing) are compensated for, or mitigated by, the statistical power gain from obtaining very large and eventually massive sample sizes. Nonetheless, the platform includes components that keep any phenotyping inaccuracy associated with the more cost-effective method of testing to a minimum. For example, to achieve standardization of the size of the stimulus across different computer displays, the platform employs a pre-test calibration module in which the user adjusts the size of a virtual object (i.e., an image of AF glasses displayed on the screen) to match a

real-world reference object (the AF glasses in the user's hands). A pre-test screen brightness calibration module is also in development. In addition, given the lack of researcher supervision for at-home testing, a brief non-rivalry physical stimulus alternation catch trial is incorporated into each rivalry test trial. If the subject responds beyond a threshold accuracy to these catch trials (for which correct perceptual durations are known), it is clear they have understood response instructions, are attending to the task, and have made few response errors. Furthermore, the platform includes a brief online assessment of visual acuity, a subject factor which can affect BRR. A subject's BR data are only included if their catch trial accuracy and visual acuity are within acceptable limits.

Despite efforts to reduce phenotyping inaccuracy, online psychophysical testing cannot control for every factor introduced by the use of at-home testing that has the potential to affect BRR. As mentioned though, the use of very large and indeed massive sample sizes, will minimize noise in the data introduced by such factors. Moreover, for studies in which a high degree of control of psychophysical factors is essential, and for which very large and massive sample sizes are not required, the online BR test can be utilized under controlled, supervised laboratory testing conditions, at least thus enabling the testing and analysis protocol standardization that has been lacking in clinical and related (e.g., personality factor) BR research. Such standardization can facilitate meaningful comparison of BR data collected from different research laboratories and will be useful to the field whether testing occurs online or in the laboratory.

While the online BR test platform will require further user-interface and technical validation testing, there appear to be no insurmountable barriers to the method facilitating genetic and clinical BR studies with large to massive sample sizes using a standardized test protocol. This will enable examination of important clinical translation questions such as whether the slow BR trait: (i) is specific to BD; (ii) provides diagnostic assistance for classical

categories in unclear psychiatric presentations of psychosis or depression; (iii) identifies biological subtypes within classical categories, (iv) identifies biological subtypes beyond classical categories (e.g., aligning with RDoC); (v) predicts risk or prognosis (including risk for developing BD in unaffected relatives of BD probands); and (vi) predicts particular treatment responsiveness irrespective of diagnosis. Initially cross-sectional studies can be performed to address such issues, followed by longitudinal prospective intervention studies assessing the efficacy of using BRR to direct intervention or treatment choices. Use of the online BR test platform can also faciliate large-scale and standardized studies of BR in psychiatric conditions other than BD, SCZ and MDD, and in developmental and neurological conditions. For example, BR studies (with small samples and non-standarized test protocols) have been performed in persons with attention deficit hyperactivity disorder[47] and autism[48-50], with a focus on BRR and MPD, respectively.

Understanding the genetic and systems level abnormalities in BD, and indeed psychiatric disorders in general, has proven more difficult than anticipated, though association signals within risk loci are being identified[22-25,51,52]. Large-scale endophenotype studies have been proposed to deal with the problem of phenotypic heterogeneity, which contributes to the limited success of psychiatric genetic studies[53]. Endophenotype BR studies may improve power of BD GWAS and shed light on the genetic and molecular basis of BD and the genetic overlap between BD and other psychiatric disorders such as SCZ[54-57]. BD GWAS using BR as an endophenotype, and BR GWAS in CTLs[33,34], would benefit from utilizing test methods that enable very large and massive sample size acquisition and standardized test protocols, such as the method discussed in the current study.

Finally, as well as potentially identifying novel pharmacotherapeutic targets, understanding genetic and molecular level processes in BD may inform understanding of the structural and systems-level abnormalities in BD, with current trends for studying the latter

also toward large-scale studies (e.g., Hibar et al.[60]). Such understanding may in turn yield new, more targeted therapeutic options. In addition, converging molecular and systems level understanding of BD (and other psychiatric disorders) may also be informed by mechanistic understanding of the endophenotype itself. In the present context, for example, a lateralized (cerebral hemisphere-based) mechanistic model of BR[21,61-63] underlies a lateralized pathophysiological model of BD[7] that: (i) has been discussed in light of robust corpus callosum abnormality findings from BD neuroimaging studies[64,65]; and (ii) has the potential for direct clinical translation using lateralized brain stimulation[7,66]. These mechanistic and pathophysiological models, and important developments in the field since their original proposal, were recently reviewed in detail[67].

Acknowledgements

We thank all volunteers who participated in the study. SMM received support from National Health and Medical Research Council, Defence Health Foundation, a 2012 NARSAD Young Investigator Grant from the Brain & Behavior Research Foundation (USA) [grant number 19163], Monash Institute of Medical Engineering, Monash University, and Victorian Government Department of Health and Human Services. TTN was supported by National Health and Medical Research Council [grant number 490976]. PCFL was supported by Victorian Government Department of Health and Human Services.

Conflict of interest

SMM is a co-inventor on lapsed University of Queensland patents related to slow BR in BD. There are no commercialization activities associated with the Binocular Rivalry Online website or test method. SMM has received equipment from Soterix Medical Inc for research. SMM has received funding from Monash Institute of Medical Engineering to examine methods for vestibular neuromodulation self-administration. If such methods are commercialised, Monash University, Monash Alfred Psychiatry Research Centre and SMM would receive royalties.

References

1. Scarr E et al. Biomarkers for psychiatry: The journey from fantasy to fact, a report of the 2013 CINP think tank. *International Journal of Neuropsychopharmacology* 2015; **18**: 1–9.
2. Malhi GS, Taniousm M, Gershon S. The lithiumeter: A measured approach. *Bipolar Disorders* 2011; **13**: 219–26.
3. Kupfer D, First M, Regier D (eds). *A research agenda for DSM-V*. Washington, DC: American Psychiatric Association; 2002.
4. Kapur S, Phillips AG, Insel TR. Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? *Molecular Psychiatry* 2012; **17**: 1174–9.
5. Miller SM et al. Slow binocular rivalry in bipolar disorder. *Psychological Medicine* 2003; **33**: 683–92.
6. Ngo TT, Mitchell PB, Martin NG, Miller SM. Psychiatric and genetic studies of binocular rivalry: An endophenotype for bipolar disorder? *Acta Neuropsychiatrica* 2011; **23**: 37–42.
7. Pettigrew JD, Miller SM. A 'sticky' interhemispheric switch in bipolar disorder? *Proceedings of the Royal Society B: Biological Sciences* 1998; **265**: 2141–8.
8. Blake R. A primer on binocular rivalry, including current controversies. *Brain and Mind* 2001; **2**: 5–38.
9. Miller SM (ed). The Netherlands: John Benjamins Publishing Company; 2013.
10. Wade NJ, Ngo TT. Early views on binocular rivalry. In Miller SM (ed). *Advances in Consciousness Research (Vol 90) The constitution of visual consciousness: Lessons from binocular rivalry*. John Benjamins Publishing Company: Amsterdam, The Netherlands, 2013. pp 77–108.
11. Blake R, Logothetis NK. Visual Competition. *Nature Reviews Neuroscience* 2002; **3**: 13–21.
12. Alais D, Blake R. *Binocular rivalry*. MIT Press: Cambridge, MA, 2005.
13. Nagamine M, Yoshino A, Miyazaki M, Takahashi Y, Nomura S. Difference in binocular rivalry rate between patients with bipolar I and bipolar II disorders. *Bipolar Disorders* 2009; **11**: 539–46.

14. Vierck E et al. Further evidence for slow binocular rivalry rate as a trait marker for bipolar disorder. *Australian and New Zealand Journal of Psychiatry* 2013; **47**: 371–9.

15. Zhu R, Ye X, Zhou X, Yang J, Yue Y, Wang K. Applications of the binocular rivalry paradigm in patients with bipolar disorder. *Chinese Journal of Nervous and Mental Diseases* 2013; **8**: 449–52.

16. Miller SM et al. Genetic contribution to individual variation in binocular rivalry rate. *Proceedings of the National Academy of Sciences of the United States of America* 2010; **107**: 2664–8.

17. Jia T et al. Difference in the binocular rivalry rate between depressive episodes and remission. *Physiology & Behavior* 2015; **151**: 272–8.

18. Law PCF, Gurvich CT, Ngo TT, Miller SM. Evidence that eye-movement profiles do not explain slow binocular rivalry rate in bipolar disorder: Support for a perceptual endophenotype. *Bipolar Disorders* 2017; **19**: 465–76.

19. Law PCF, Paton BK, Riddiford JA, Gurvich CT, Ngo TT, Miller SM. No relationship between binocular rivalry rate and eye-movement profiles in healthy individuals: A Bayes factor analysis. *Perception* 2015; **44**: 643–61.

20. Shannon RW, Patrick CJ, Jiang Y, Bernat E, He S. Genes contribute to the switching dynamics of bistable perception. *Journal of Vision* 2011; **11**: 1–7.

21. Ngo TT, Barsdell WN, Law PCF, Miller SM. Binocular rivalry, brain stimulation and bipolar disorder. In Miller SM (ed). *Advances in Consciousness Research (Vol 90) The constitution of visual consciousness: Lessons from binocular rivalry*. John Benjamins Publishing Company: Amsterdam, The Netherlands, 2013. pp 211–52.

22. Hou L et al. Genome-wide association study of 40,000 individuals identifies two novel loci associated with bipolar disorder. *Human Molecular Genetics* 2016; **25**: 3383–94.

23. Mullins N et al. (2021) Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology. *Nature Genetics.* 2021; **53**, 817–29.

24. Maier RM, Visscher PM, Robinson MR, Wray NR. Embracing polygenicity: A review of methods and tools for psychiatric genetics research. *Psychological Medicine* 2017; **48**: 1055–

67.

25. Ivleva EI, Morris DW, Moates AF, Suppes T, Thaker GK, Tamminga CA. Genetics and intermediate phenotypes of the schizophrenia—bipolar disorder boundary. *Neuroscience and Biobehavioral Reviews* 2010; **34**: 897–921.

26. Law PCF, Paton BK, Thomson RH, Liu GB, Miller SM, Ngo TT. Dichoptic viewing methods for binocular rivalry research: Prospects for large-scale clinical and genetic studies. *Twin Research and Human Genetics* 2013; **16**: 1033–78.

27. Conus P, McGorry PD. First-episode mania: A neglected priority for early intervention. *Australian and New Zealand Journal of Psychiatry* 2002; **36**: 158–72.

28. Joyce PR. Age of onset in bipolar affective disorder and misdiagnosis as schizophrenia. *Psychological Medicine* 1984; **14**: 145–9.

29. Goodwin FK, Jamison KR, Ghaemi SN. *Manic-depressive illness: bipolar disorders and recurrent depression.* Oxford University Press: New York, 2007.

30. Nasrallah HA. Consequences of misdiagnosis: inaccurate treatment and poor patient outcomes in bipolar disorder. *Journal of Clinical Psychiatry* 2015; **76**: e1328.

31. Hirschfeld RM, Lewis L, Vornik LA. Perceptions and impact of bipolar disorder: how far have we really come? Results of the national depressive and manic-depressive association 2000 survey of individuals with bipolar disorder. *Journal of Clinical Psychiatry* 2003; **64**: 161–74.

32. Xiao G et al. Slow binocular rivalry as a potential endophenotype of schizophrenia. *Frontiers in Neuroscience* 2018; **12**: 1–9.

33. Bosten JM, Goodbourn PT, Lawrance-Owen AJ, Bagary G, Hogg RE, Mollon JD. A population study of binocular function. *Vision Research* 2015; **110**: 34–50.

34. Chen B et al. Genomic analyses of visual cognition: Perceptual rivalry and top-down control. *Journal of Neuroscience* 2018; **38**: 9668–78.

35. Wade NJ. Monocular and binocular rivalry between contours. *Perception* 1975; **4**: 85–95.

36. Thomas J. Binocular rivalry: The effects of orientation and pattern color arrangement. *Perception & Psychophysics* 1978; **23**: 360–2.

37. Kitterle FL, Thomas J. The effects of spatial frequency, orientation, and color upon binocular

rivalry and monocular pattern alternation. *Bulletin of the Psychonomic Society* 1980; **16**: 405–7.

38. Levelt W. *On binocular rivalry*. Institute for Perception RVO-TNO: Soesterberg, The Netherlands, 1965.

39. Law PCF, Miller SM, Ngo TT. The effect of stimulus strength on binocular rivalry rate in healthy individuals: Implications for genetic, clinical and individual differences studies. *Physiology & Behavior* 2017; **181**: 127–36.

40. Sheehan DV et al. The Mini International Neuropsychiatric Interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview. *Journal of Clinical Psychiatry* 1998; **59**: 22–33.

41. Anderson EC et al. Smiles may go unseen in generalized social anxiety disorder: Evidence from binocular rivalry for reduced visual consciousness of positive facial expressions. *Journal of Anxiety Disorders* 2013; **27**: 619–26.

42. Carter OL et al. Modulating the rate and rhythmicity of perceptual rivalry alternations with the mixed 5-HT2A and 5-HT1A agonist psilocybin. *Neuropsychopharmacology* 2005; **30**: 1154–62.

43. Nagamine M, Yoshino A, Miyazaki M, Takahashi Y, Nomura S. Effects of selective 5-HT1A agonist tandospirone on the rate and rhythmicity of binocular rivalry. *Psychopharmacology* 2008; **198**: 279–86.

44. van Loon AM, Knapen T, Scholte HS, St. John-Saaltink E, Donner TH, Lamme VAF. GABA shapes the dynamics of bistable perception. *Current Biology* 2013; **23**: 823–7.

45. Ukai K, Ando H, Kuze J. Binocular rivalry alternation rate declines with age. *Perceptual Motor Skills* 2003; **97**: 393–7.

46. Nagamine M et al. Accelerated binocular rivalry with anxious personality. *Physiology & Behavior* 2007; **91**: 161–5.

47. Amador-Campos JA, Aznar-Casanova JA, Moreno-Sanchez M, Medina-Pena A, Ortiz-Guerra JJ. Psychometric properties of a test for ADHD based on binocular rivalry. *Spanish Journal of Psychology* 2013; **16**: 1–8.

48. Freyberg J, Robertson CE, Baron-Cohen S. Reduced perceptual exclusivity during object and grating rivalry in autism. *Journal of Vision* 2015; **15**: 1–12.

49. Robertson CE, Kravitz DJ, Freyberg J, Baron-Cohen S, Baker CI. Slower rate of binocular rivalry in autism. *Journal of Neuroscience* 2013; **33**: 16983–91.

50. Said CP, Egan RD, Minshew NJ, Behrmann M, Heeger DJ. Normal binocular rivalry in autism: Implications for the excitation/inhibition imbalance hypothesis. *Vision Research* 2013; **77**: 59–66.

51. Chen DT et al. Genome-wide association study meta-analysis of European and Asian-ancestry samples identifies three novel loci associated with bipolar disorder. *Molecular Psychiatry* 2013; **18**: 195–205.

52. Mühleisen TW et al. Genome-wide association study reveals two new risk loci for bipolar disorder. *Nature Communications* 2014; **5**: 1–8.

53. Hasler G, Drevets WC, Gould TD, Gottesman II, Manji HK. Toward constructing an endophenotype strategy for bipolar disorders. *Biological Psychiatry* 2006; **60**: 93–105.

54. Forstner AJ et al. Identification of shared risk loci and pathways for bipolar disorder and schizophrenia. *PLoS One* 2017; **12**: 1–14.

55. Lee SH et al. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nature Genetics* 2013; **45**: 984–94.

56. Ruderfer DM et al. Genomic dissection of bipolar disorder and schizophrenia, including 28 subphenotypes. *Cell* 2018; **173**: 1705–15.

57. Tao R et al. Transcript-specific associations of SLC12A5 (KCC2) in human prefrontal cortex with development, schizophrenia, and affective disorders. *Journal of Neuroscience* 2012; **32**: 5216–22.

58. Rizak J, Tan H, Zhu H, Wang JF. Chronic treatment with the moodstabilizing drug lithium up-regulates nuclear factor E2-related factor 2 in rat pheochromocytoma PC12 cells in vitro. *Neuroscience* 2014; **256**: 223–9.

59. Beasley CL, Pennington K, Behan A, Wait R, Dunn MJ, Cotter D. Proteomic analysis of the anterior cingulate cortex in the major psychiatric disorders: Evidence for disease-associated

changes. *Proteomics* 2006; **6**: 3414 –25.

60. Hibar DP et al. Cortical abnormalities in bipolar disorder: An MRI analysis of 6503 individuals from the ENIGMA Bipolar Disorder Working Group. *Molecular Psychiatry* 2017; **23**: 932–42.
61. Miller SM. Binocular rivalry and the cerebral hemispheres with a note on the correlates and constitution of visual consciousness. *Brain and Mind* 2001; **2**: 119–49.
62. Miller SM et al. Interhemispheric switching mediates perceptual rivalry. *Current Biology* 2000; **10**: 383–92.
63. Pettigrew JD. Searching for the switch: Neural bases for perceptual rivalry alternations. *Brain and Mind* 2001; **2**: 85–118.
64. Cullen KR, Lim KO. Toward understanding the functional relevance of white matter deficits in bipolar disorder. *JAMA Psychiatry* 2014; **71**: 362–4.
65. Sarrazin S et al. A multicenter tractography study of deep white matter tracts in bipolar I disorder: Psychotic features and interhemispheric disconnectivity. *JAMA Psychiatry* 2014; **71**: 388–96.
66. Miller SM. Vestibular neuromodulation: Stimulating the neural crossroads of psychiatric illness. *Bipolar Disorders* 2016; **18**: 539–43.
67. Miller SM. Fluctuations of consciousness, mood, and science: The interhemispheric switch and sticky switch models two decades on. *Journal of Comparative Neurology* 2020; **528**; 3171–97.
68. Hollins M. The effect of contrast on the completeness of binocular rivalry suppression. *Perception & Psychophysics* 1980; **27**: 550–6.
69. Suzuki S, Grabowecky M. Long-term speeding in perceptual switches mediated by attention-dependent plasticity in cortical visual processing. *Neuron* 2007; **56**: 741–53.

TABLES

TABLE 1. Average and retest reliability of binocular rivalry rates across stimulus conditions.

	Binocular rivalry rate					Retest reliability of binocular rivalry rate			
	Persons with BD		CTLs			Persons with BD		CTLs	
	Median ± *MAD*	Mean ± *SD*	Median ± *MAD*	Mean ± *SD*	*p*-value [a]	*ρ(CI)*	*p*-value [a]	*ρ(CI)*	*p*-value [b]
Stationary green-PF gratings [b]	0.27 ± 0.09	0.32 ± 0.18	0.38 ± 0.08	0.41 ± 0.15	4.91×10^{-2}	—	—	—	—
4 c/s green-PF gratings	0.26 ± 0.11	0.30 ± 0.16	0.50 ± 0.14	0.51 ± 0.21	1.16×10^{-3}	0.77(0.50–0.90)	4.20×10^{-5}	0.83(0.61–0.93)	3.70×10^{-6}
8 c/s green-PF gratings	0.28 ± 0.05	0.32 ± 0.18	0.49 ± 0.15	0.56 ± 0.29	2.60×10^{-4}	0.46(0.02–0.75)	2.02×10^{-2}	0.83(0.61–0.93)	2.76×10^{-6}
4 c/s red/blue-AF gratings	0.34 ± 0.07	0.35 ± 0.14	0.47 ± 0.11	0.54 ± 0.22	1.16×10^{-3}	0.79(0.53–0.91)	1.65×10^{-5}	0.71(0.39–0.88)	2.02×10^{-4}
8 c/s red/blue-AF gratings	0.32 ± 0.06	0.35 ± 0.15	0.50 ± 0.16	0.51 ± 0.20	1.22×10^{-2}	0.73(0.43–0.89)	1.15×10^{-4}	0.76(0.48–0.90)	4.66×10^{-5}
4 c/s red/blue-AF rings	0.25 ± 0.06	0.25 ± 0.12	0.40 ± 0.13	0.43 ± 0.17	1.44×10^{-3}	0.48(0.05–0.76)	1.71×10^{-2}	0.75(0.46–0.90)	6.69×10^{-5}
8 c/s red/blue-AF rings	0.24 ± 0.08	0.24 ± 0.10	0.42 ± 0.10	0.45 ± 0.18	4.51×10^{-4}	0.25(-0.22–0.62)	0.15	0.77(0.50–0.90)	3.23×10^{-5}

BD: bipolar disorder. CTL: control. PF: polarization filter. AF: anaglyph filter. c/s: cycles/s. *SD*: standard deviation. *MAD*: median absolute deviation. *ρ*: Spearman's ρ. *CI*: 95% confidence interval.
[a] Bonferroni-adjusted α: 0.05/7 Mann-Whitney U tests $=7.14\times10^{-3}$. [b] 0 cycles/s. [b] One tailed. Bonferroni-adjusted α: 0.05/6 Spearman's ρ tests$=8.33\times10^{-3}$. [c] Two tailed. Bonferroni-adjusted α: 0.05/6 Spearman's ρ tests$=8.33\times10^{-3}$.

TABLE 2. Average and retest reliability of mixed-percept durations across stimulus conditions.

	Mixed-percept duration					Retest reliability of mixed-percept duration			
	Persons with BD		CTLs			Persons with BD		CTLs	
	Median ± *MAD*	Mean ± *SD*	Median ± *MAD*	Mean ± *SD*	*p*-value [a]	*ρ(CI)*	*p*-value [a]	*ρ(CI)*	*p*-value [c]
Stationary green-PF gratings [b]	114.36 ± 62.98	106.12 ± 76.64	64.61 ± 32.07	66.53 ± 50.72	7.63×10^{-2}	—	—	—	—
4 c/s green-PF gratings	69.01 ± 18.70	75.95 ± 48.77	43.72 ± 22.46	47.59 ± 31.35	3.04×10^{-2}	0.64(0.28–0.84)	2.27×10^{-3}	0.65(0.29–0.85)	1.75×10^{-3}
8 c/s green-PF gratings	104.63 ± 38.96	107.38 ± 107.38	54.51 ± 22.14	63.65 ± 36.76	4.68×10^{-3}	0.72(0.41–0.88)	3.56×10^{-4}	0.58(0.19–0.81)	7.55×10^{-3}
4 c/s red/blue-AF gratings	38.73 ± 23.77	46.38 ± 42.43	19.15 ± 14.10	28.17 ± 32.28	0.13	0.54(0.13–0.79)	1.41×10^{-2}	0.61(0.23–0.83)	4.09×10^{-3}
8 c/s red/blue-AF gratings	44.22 ± 27.53	52.54 ± 48.44	36.13 ± 22.31	35.18 ± 28.68	0.37	0.68(0.34–0.86)	1.10×10^{-3}	0.58(0.19–081)	7.17×10^{-3}
4 c/s red/blue-AF rings	45.28 ± 27.22	49.88 ± 49.46	13.71 ± 13.71	45.77 ± 65.24	0.38	0.49(0.06–0.77)	2.88×10^{-2}	0.81(0.57–0.92)	1.30×10^{-5}
8 c/s red/blue-AF rings	35.96 ± 27.61	50.18 ± 49.61	17.40 ± 17.40	38.37 ± 50.34	0.24	0.62(0.25–0.83)	3.39×10^{-3}	0.76(0.48–0.90)	9.51×10^{-5}

BD: bipolar disorder. CTL: control. PF: polarization filter. AF: anaglyph filter. c/s: cycles/s. *SD*: standard deviation. *MAD*: median absolute deviation. *ρ*: Spearman's ρ. *CI*: 95% confidence interval.
[a] Bonferroni-adjusted α: 0.05/7 Mann-Whitney U tests $=7.14\times10^{-3}$. [b] 0 cycles/s. [b] One tailed. Bonferroni-adjusted α: 0.05/6 Spearman's ρ tests=8.33×10^{-3}. [c] Two tailed. Bonferroni-adjusted α: 0.05/6 Spearman's ρ tests=8.33×10^{-3}.

TABLE 3. Comparison of stimulus conditions assessed according to key criteria.

	Can be used for at-home BR testing	Binocular rivalry rate [a]						
		Slower in BD than in CTLs	Similar to that elicited with same-speed green-PF gratings		Correlation with same-speed green-PF gratings [b]		Between-session reliability	
			BD	CTLs	BD	CTLs	BD	CTLs
Polarization filter (PF) method								
Stationary green gratings	NO	NO	—	—	—	—	—	—
4 c/s green gratings	NO	YES	—	—	—	—	HIGH	HIGH
8 c/s green gratings	NO	YES	—	—	—	—	*n.s.*	HIGH
Anaglyph filter (AF) method								
4 c/s red/blue gratings	YES	YES	YES	YES	MODERATE	HIGH	HIGH	MODERATE
8 c/s red/blue gratings	YES	NO	YES	YES	*n.s.*	HIGH	MODERATE	HIGH
4 c/s red/blue rings	YES	YES	YES	YES	MODERATE	HIGH	*n.s.*	HIGH
8 c/s red/blue rings	YES	YES	YES	YES	*n.s.*	MODERATE	*n.s.*	HIGH

BR: binocular rivalry. BD: bipolar disorder. CTL: control. c/s: cycles/second. *n.s.*: not significant.

[a] Mixed-percept duration (MPD) considerations are also relevant to stimulus optimization however MPD data are not shown here due to caution required in their interpretation (see Discussion). [b] Positive correlation.

FIGURE LEGENDS

FIGURE 1. Binocular rivalry (BR) and experiment protocol. **(A)** Presenting dissimilar images simultaneously, one to each eye, causes each image to stochastically alternate in perception. Mixed or piecemeal percepts occur occasionally during the transition between perception of the presented images. Images are presented on a lower luminance uniform black background. Arrows adjacent to the presented stimuli denote the direction of grating drift. **(B)** All subjects completed three testing sessions of BR viewing, each on a separate day, with each session involving multiple blocks each with four trials. Stimulus type and drift speed (4 and 8 cycles/s) were counterbalanced across case-control pairs. The full protocol description has been published previously (Law et al.[18]). Vertical gratings drifted rightward and horizontal gratings downward (i.e., in orthogonal directions), while red rings expanded and blue rings contracted (i.e., in opposite directions). Dashed boxes around the stimuli denote that the enclosed test blocks were counterbalanced across case-control pairs. The first block in each testing session served to adequately stabilize BRRs for the remaining test blocks (discussed in Miller et al.[5]; see also Hollins[68]; Suzuki & Grabowecky[69]) and was excluded from analysis. **(C)** The retest session included five blocks retesting the participant's drifting grating stimulus protocol as viewed in the first testing session, followed by **(D)** exploratory test blocks with red rings expanding and blue rings contracting at either 4 or 8 cycles/s viewed through AFs (reported in the current paper and retested in the third testing session) and coherence rivalry stimuli viewed through PFs (not depicted and the subject of a separate report).

FIGURES

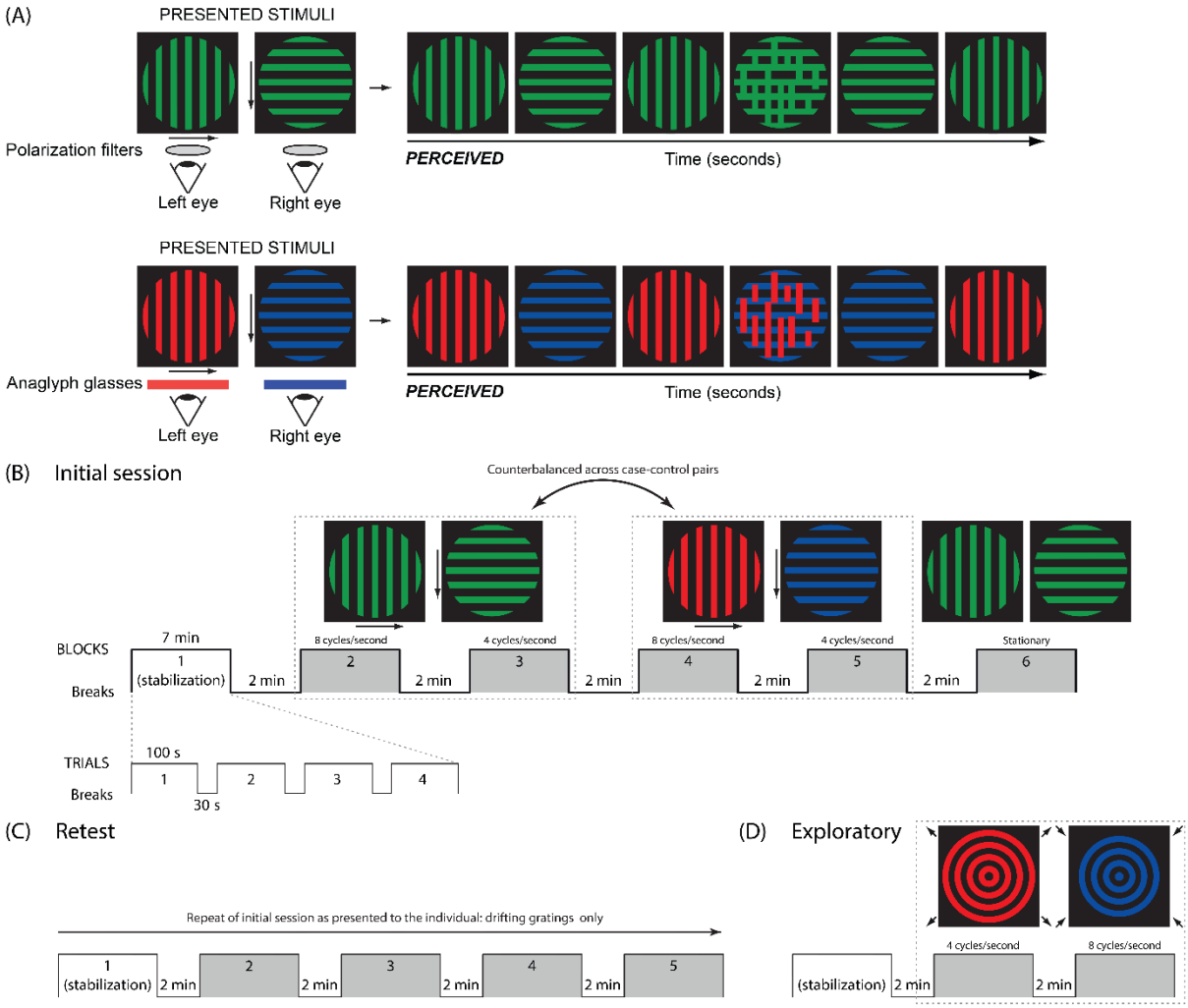

FIGURE 1.