



# Assessing the utility of a virtual-reality neuropsychological test battery, 'CONVIRT', in detecting alcohol-induced cognitive impairment

Isabella Amato<sup>1</sup> · Aleshia Nanev<sup>1</sup> · Stefan Piantella<sup>1</sup> · Kira-Elise Wilson<sup>1</sup> · Rowena Bicknell<sup>1</sup> · Rachael Heckenberg<sup>1</sup> · Ben Horan<sup>2</sup> · Paul Maruff<sup>3</sup> · Bradley Wright<sup>1</sup>

Accepted: 10 September 2020  
© The Psychonomic Society, Inc. 2020

## Abstract

New technologies such as virtual reality (VR) and eye-tracking software have paved the way for more sophisticated and ecologically valid measures of cognitive function. Testing the sensitivity and reliability of such measurements in response to acute alcohol intoxication provides a first step in establishing how these measures may operate in relation to cognitive impairments observed post-concussion. Healthy young adults ( $N = 54$ ,  $M = 20.65$ ,  $SD = 2.06$ , 30 females) completed the CONVIRT test battery (manual simple and choice reaction-time and saccade reaction-time) at three breath alcohol concentration (BrAC) levels: 0.00%<sup>T1</sup>, 0.05%<sup>T2</sup>, 0.08%<sup>T3</sup>. Participants consumed alcoholic beverages at 30-min intervals, with BrAC monitored at 15-min intervals using a breathalyser. All three CONVIRT measures were sensitive to changes in cognitive performance induced by alcohol at BrAC levels at or exceeding 0.05%. A composite measure was also sensitive to alcohol intoxication (Cohen's  $d = .85$  at BrAC = 0.05%;  $d = 1.20$  at BrAC = 0.08%). Strong test–retest reliability was observed (all  $r < .80$ ), with no gender differences noted. CONVIRT measures were reliable and detected dose-dependent changes in alcohol-induced cognitive impairment. Potentially, the ecologically valid measures may assist in better quantifying the effects of conditions such as concussion, on cognitive performance.

**Keywords** Decision-making · Attention · Saccade · Concussion · Jockeys · Virtual reality

## Introduction

The field of neuropsychology is scaffolding on the solid foundations of validated assessment models and beginning to embrace new technology for clinical decision-making. Specifically, as virtual-reality (VR) technology and environments confer several advantages over traditional neuropsychological assessments, there has been increasing interest in the use of virtual-reality contexts for cognitive testing to inform neuropsychological decision-making (e.g., Howett et al., 2019; Wiener et al., 2019). VR environments are immersive

and able to consistently hold distractors such as light, sound, and movement constant for each test occasion while providing a closer approximation of the potentially dynamic and noisy environments in which cognitive demands are made. Further, head-mounted displays (HMDs) have recently emerged, providing the ability to readily immerse users in high-quality virtual environments (Salomoni et al., 2017), with eye-tracking technology integrated into these headsets. Given these advances, there are now new opportunities to refine assessments of CNS disruption using such technology.

The CONVIRT battery was developed within a VR framework to assess and determine if the cognitive abilities of a concussed horse jockey had returned to pre-concussion levels. The CONVIRT battery uses manual simple and choice reaction time, and saccadic reaction time to assess components of attention, decision-making, and visual processing speed within a dynamic VR environment where these cognitive operations are assessed in the context of a virtual horse race (see Horan et al., 2020 for a full description). The greatest threats to the health of professional jockeys are concussion and head injury. Jockeys have the highest rates of fatality (per minute

---

✉ Bradley Wright  
b.wright@latrobe.edu.au

<sup>1</sup> Bundoora Campus, Psychology and Counselling, La Trobe University, Bundoora, Australia

<sup>2</sup> Faculty of Science and Technology, School of Engineering, Deakin University, Melbourne, Australia

<sup>3</sup> CogState Ltd, Melbourne, Australia

of participation) and concussion of all sportspersons (Turner, McCrory, & Halley, 2002). In this context, cognitive assessments form an important part of concussion management systems that are used to determine when a jockey is fit-to-ride post-concussion and has serious implications for the jockey, other jockeys and the medical personnel responsible for such decisions.

In a large study comparing the CONVIRT battery with validated and comparable tests from the computerised Cogstate battery, ( $N = 165$ ), the CONVIRT battery of virtual-reality cognitive tests showed excellent test–retest reliability, and convergent validity with the Cogstate attention and decision-making tasks (identification and detection tasks; Horan et al., 2020). Evidence of greater physiological arousal during performance from measures of heart rate and heart rate variability also suggested that performance on the CONVIRT battery was more ecologically valid than the Cogstate battery. These results indicate the potential utility of the CONVIRT VR system for concussion management programs for jockeys. However, because the classification of concussion often relies on decisions about whether cognition has deteriorated following injury, before the utility of the system can be investigated in clinical use a thorough psychometric evaluation of the CONVIRT system and its measurements of responsiveness are required—that is, to determine the sensitivity of the CONVIRT system to true cognitive change. It is difficult to determine the sensitivity of new tests to cognitive change because it requires experimental contexts where a true and meaningful (and reversible) cognitive change can be applied and measured.

One model for determining the sensitivity of cognitive instruments to true cognitive change is to borrow from psychopharmacological models in which cognition is impaired temporarily by administration of a CNS active sedative drug. The dose-related effects of alcohol on cognition are well known, and acute low-dose alcohol administration models have been employed to quantify the sensitivity of such cognitive assessments (e.g., Falleti et al., 2003; Maruff et al., 2005). As drink-driving legislation and enforcement has led to widespread societal knowledge about the adverse effects of alcohol on purposive behaviour, the importance of cognitive impairment can be expressed within the same framework. With respect to concussion-related cognitive impairment comparative analyses revealed that cognitive decline in concussed Australian Rules Footballers 2–3 days post-concussion was similar to that observed at a blood alcohol concentration (BAC) of 0.10% (Maruff et al., 2001). The standard BAC levels used internationally for drink-driving legislation are 0.05% and 0.08%, as it is accepted that these nominated BAC levels present an unacceptably high risk for driving. The 0.08% intoxication level has been associated with very large decrements in cognitive performance on the Cogstate battery

of computerised neuropsychological tests (e.g., Cohen's  $d > 1.00$ , Maruff et al., 2005), and a number of recent studies (e.g., Cash, Peacock, Barrington, Sinnott, & Bruno, 2015; Dry, Burns, Nettlebeck, Farquharson, & White, 2012; Jongen, Vuurman, Ramaekers, & Vermeeren, 2016) have sought to determine the sensitivity to important cognitive change using this alcohol challenge as a setting.

One aspect of the CONVIRT system that is novel in the area of concussion management is the assessment of saccadic eye movements. Saccades are a type of ocular movement constituting rapid, ballistic eye movements that shift the gaze, either voluntarily or reflexively, to new areas of interest and help bring a target into focus (Ramat, Leigh, Zee, & Optican, 2007). Research has shown delays in saccade speed post-concussion (see Sneigireva et al., 2018 for a review) that can be evident up to three weeks post-injury (Mullen et al., 2014). Furthermore, experimental psychopharmacological models show longer latency to initiate a saccade and slower saccade acceleration in response to relatively low doses of alcohol (Vassallo & Abel, 2002; Blekher et al., 2002a; King & Byars, 2004; Vorstius et al., 2008; Wegner and Fahle 1999, Roche & King, 2010). Therefore, measurement of saccades should also provide a sensitive index of CNS disruption in concussion and consequently it is important to understand the sensitivity of the operationalization of this measurement within the CONVIRT system.

In the present study, we used an acute administration of alcohol model as a means of inducing temporary cognitive impairment in a group of healthy young adults. We used this to test the CONVIRT system's sensitivity to change. Given the need for the measures to be administered repeatedly, we assessed the test–retest reliability of each measure and anticipated satisfactory reliability on each of the tests as the BrAC was consistently applied across all participants. The second hypothesis was that low-level alcohol intoxication at levels classified as unacceptable for driving a motor vehicle (i.e., BAC 0.08%) would result in performance declines from baseline for all CONVIRT measures.

## Materials and methods

### Participants

Participants were healthy young adults ( $N = 54$ , females = 30), aged 18 to 28 years ( $M = 20.65$ ,  $SD = 2.06$ ). Using a script, participants were recruited face-to-face at La Trobe University, Australia. Participants were included if they were currently full-time students (a criterion for a separate study) and could read English, and excluded if they reported being ill on the day of testing, had never consumed alcohol, regularly used drugs or alcohol, or received a score above 8 on the

Alcohol Use Disorders Identification Test (AUDIT; Saunders, Aasland, Babor, de la Fuente, & Grant, 1993). Participants were also excluded if they had a history of head injury, psychiatric/neurological illness, sleep or emotional disorders or used medications that altered cognitive functioning. Participants provided written informed consent in line with institutional ethics (HEC19036) and were compensated for their time with a \$100 (AUD) shopping voucher.

## Materials

**Alcohol consumption screening** The ten-item Alcohol Use Disorders Identification Test (AUDIT; Saunders et al., 1993) was used to assess problematic drinking behaviours to screen for participants who may be alcohol-dependent. Individuals scoring higher than an 8 on the AUDIT were excluded from the study. The questionnaire assesses alcohol consumption (three items), drinking behaviour (three items) and consequences related to drinking (four items). Items are scored according to frequency (1 – never, to 5 – daily/almost daily) and items nine and ten are rated on a 3-point Likert scale (1 – no, to 3 – yes, during the last year) where participants specify their level of agreement to statements. An AUDIT score of eight or above indicated risky/hazardous levels of alcohol use. The AUDIT has excellent internal consistency ( $r = .81$ ; Shields & Caruso, 2003) and strong test–retest reliability (Selin, 2003).

**BrAC calculator** An online calculator was used to estimate the amount of alcohol needed for each individual to reach a BrAC of 0.05% and 0.08% (<https://www.rupissed.com/>). This is a widely used Australian website that predicts BRAC levels based on demographics (i.e., height, weight, age, ethnicity, and gender), alcohol type and quantity, and rate of consumption.

**Breathalyser** The Alcolizer LE5 breathalyser (2016; Alcolizer Pty Ltd., Brisbane, Queensland, Australia; model number: 400220CS) was used to monitor the BRAC of participants throughout the testing day. The LE5 breathalyser is certified by Australian standards as a police-grade device that provides an indirect, approximate BAC measurement (Sorbello, Devilly, Allen, Hughes, & Brown, 2018). While blood sampling collects true BAC, breathalysers are non-invasive, reduce administration time, and provide immediate results (Smith, 2011). They are also an accurate and reliable measure when compared to blood sampling in both drink-driving data ( $r = .94$ ; Zuba, 2008) and laboratory studies ( $r = .88$ ; Schechtman & Shinar, 2011). The Alcolizer LE5 breathalyser has excellent test–retest reliability ( $r = .99$ ), inter-instrument reliability and convergent validity ( $r = .93$  to  $r = .97$ ; Sorbello et al., 2018) with blood tests.

## The CONVIRT Tool

The CONVIRT battery operates through a Gigabyte P35 laptop, and comprises the CONVIRT VR application, FOVE 0 Eye Tracking VR Head Mounted Display (HMD), and customised wireless riding crop with push button (Figure 1). The FOVE 0 HMD contains built-in eye-tracking functionality and displays a dynamic virtual environment to the participant while tracking gaze direction and head rotation. The eye tracking unit embedded in the FOVE HMD has a tracking accuracy of less than 1 degree and a refresh rate of 120 Hz (<https://www.getfove.com>). Through the HMD, participants experience the first-person perspective of a jockey riding a horse in a horserace. Full specifications of the Gigabyte P35 laptop, FOVE 0 HMD, and CONVIRT VR application are reported previously (Horan et al., 2020).

The CONVIRT battery comprises three computerised cognitive tests: the saccadic reaction time test (SAC-VR), the detection test (DET-VR), and the identification test (IDN-VR). In each test, participants respond to target shapes appearing in the virtual scene. Prior to each test, instructions are presented in the visual display, and participants undertake a short practice trial. The test-battery colour scheme is suitable for individuals who are colour-blind.

**Visual processing speed** The first test in the CONVIRT battery is the saccadic reaction time test (SAC-VR), which measures visual processing speed. As part of the eye-tracking functionality, the participant's eye movements are displayed in the virtual world by two small eye indicator shapes (one for each eye). The eye indicator shapes move across the environment in concert with the eye movement of the user. Participants start by fixating their gaze on a grey sphere for 2 s (the green eye



**Fig. 1** The CONVIRT test battery set-up, comprising the CONVIRT VR application, FOVE 0 Eye-Tracking VR Head-Mounted Display, riding crop with push button, running through a Gigabyte P35 laptop

shapes converge on the grey sphere), presented in the centre of an invisible 180-degree arc. The arc is positioned on an invisible vertical plane situated in front of the user. Following this, the grey sphere disappears, and a blue sphere appears at a random position on the arc (Fig. 2). The participant must move their gaze, as quickly as possible, towards the blue sphere, which will then ‘explode’ in response (with an explosion sound-effect and animated disintegration of the sphere). The grey sphere then reappears, and the process is repeated until the participant has fixated their eyes and ‘exploded’ 35 blue spheres.

The SAC-VR measurement is taken as the time for the participant’s gaze to reach 50% of the distance towards the blue sphere. This measures the time taken, or response latency, to initiate and begin the saccade (Orban De Xivry & Lefèvre, 2007). This approach minimises the potential of a false-positive influence from an aberrant saccade. The second 50% of the distance towards the target was not used because this component of ocular movement incorporates a range of more complex neural processes, such as deceleration of the saccade to improve accuracy when converging on the target (Orban De Xivry & Lefèvre, 2007).

**Attention** Following the SAC-VR test, participants undergo the detection test (DET-VR), which measures simple reaction time. In this test, participants are presented with an orange triangle, appearing at a random point on the 180-degree arc. After seeing a triangle, the participant must respond as quickly as possible by pressing the push button on the customised riding crop. The triangle will then disappear for a duration of between one and 2.37 s before another orange triangle appears at a random point on the arc. The process repeats for a duration of 120 s and 35 triangles are presented in total. Simple reaction time is measured based on the time elapsed (in milliseconds) between each triangle appearing and the riding crop push button being pressed. False positives occurred when the

button was pressed with no triangle present and were also recorded. Tests with false positives constituting more than 10% of all button presses were rendered invalid. No tests were excluded based on this criterion.

**Decision-making** The final CONVIRT test, the identification test (IDN-VR) is a measure of choice reaction time. Participants are presented with a random shape (orange or blue, sphere or triangle) at a random point on the 180-degree arc. When seeing an orange sphere, participant’s need to respond by pressing the riding crop button as quickly as possible. They must refrain from pushing the button in response to other stimuli. The test runs for a duration of 120 s, with 31 shapes presented in total. Choice reaction time is taken as the latency (in milliseconds) between an orange sphere appearing and the push button being pressed. False positives were recorded when the button was pressed in response to a blue sphere, or blue or orange triangle, or before the orange sphere had appeared. Tests with false positives constituting more than 20% of all button presses were rendered invalid. No tests were excluded based on this criterion.

**Composite score** In line with our earlier work (Horan et al., 2020), a composite score (COMP) was computed for cognitive performance on the three CONVIRT tests (i.e., SAC-VR, DET-VR, IDN-VR). The COMP was computed for each individual by averaging the standardised score of the three tests. This allowed for an assessment of change across differing BrAC levels within-individuals. The COMP was utilised to measure the reliability of the three tests. As standardised scores have a mean of zero, it was not possible to use COMP to compare the phases of differing BrACs. To counter this, the scores of each subtest were summed and the COMP formula was  $COMP = OMS + CRT + SRT$ . This method corresponded with  $r$  values of .99, .99, and .99, when compared to each test with the composite measure calculated with standard scores.



**Fig. 2** The virtual display of the saccadic reaction time test (SAC-VR) during the practice trial. Note. ‘A’ represents the participant focussing their eyes on the grey sphere to begin, ‘B’ looking at the blue target

sphere, ‘C’ the exploded blue sphere, and ‘D’ the user focusing on the grey sphere again before the next target appears randomly

## Procedure

Participants were screened using the AUDIT and provided basic demographic information (i.e., age, gender, height, and weight), prior to taking part in the study. Participants were excluded if they received an AUDIT score of above 8, however, no participants were excluded based on this criterion. The necessary information was entered into the BRAC calculator to determine the approximate amount of standard alcoholic drinks needed for individual participants to reach BrACs of 0.05% and 0.08%.

Two researchers were present during testing, wearing white lab coats. Two participants were tested in parallel in separate rooms, at the same time, with approximately four participants tested over a single day. This study is a part of a larger research project assessing the impact of genotype and alcohol consumption upon measures of autonomic reactivity, but these data are not presented here.

After providing informed consent, participants were instructed to fit a heart rate (HRV) monitor and electrodes, which was worn for the duration of testing. Participants also provided two saliva samples: one at baseline, and one following completion of the CONVIRT tests. These results are not reported in the present study.

After providing the baseline saliva sample, participants were then fitted with the CONVIRT headset and given the opportunity to become familiar with the riding crop and virtual reality environment. The CONVIRT subtests were presented in the same order for all participants (i.e., SAC-VR, DET-VR, IDN-VR). For each test, participants were provided with instructions and given a practice trial to ensure they were familiar with the assessment requirements and the use of equipment before they completed the test. Participants completed the CONVIRT battery across three trials with differing BrAC levels: baseline 0.00%, 0.05%, and 0.08% (Table 1).

Following the baseline phase participants consumed an alcoholic beverage (60% mixer, 40% vodka) at an individual dose until the desired BrAC level was reached. Participants' BrAC levels were monitored using the LE5 Alcolizer breathalyser at 15-min intervals following alcohol consumption. To determine the booster dose required to reach 0.08%, we manipulated the BrAC calculator by using the necessary

**Table 1.** Order of phases of alcohol ingestion and timeline for CONVIRT testing

PHASE	BrAC = 0.00%	BrAC = 0.05%	BrAC = 0.08%
<b>TIME</b>	<i>M</i> 0 min. → 28 min.	→ 2h 32 min.	→ 3h 54 min.
<b>ELAPSED</b>	<i>SD</i> 5 min	50 min	49 min

*Note.* Time elapsed refers to the time CONVIRT testing was completed in each phase. Assessment of each of the three CONVIRT subtests took approximately 10 min to complete

demographic data, and if required, adjusted up or down, how many drinks had been consumed in the timeframe to provide the current BrAC % of the participant. We then used the calculator to determine if a 15-ml or 30-ml dose was required to reach 0.08%.

After completing the final CONVIRT test, participants were given the opportunity to eat and watch television while their BrAC returned to 0.00%. Participants' BrAC was monitored every 20 min until a reading of 0.00% was observed, at which point participants could leave. If a participant's BrAC did not reach 0.00% by 5:00 pm, they were provided with a taxi voucher to ensure their safe arrival home.

## Data analysis

To assess the reliability of the CONVIRT battery across the three trials (i.e., BrAC% of 0.00, 0.05, 0.08), an intra-class correlation coefficient was calculated. Then, to assess if the CONVIRT battery was a sensitive detector of cognitive impairment induced by alcohol consumption, a series of one-way repeated measures ANOVAs were conducted with appropriate alpha-corrected post hoc analyses of the pairwise comparisons.

## Results

### Data management

As is common in measures of reaction time speed (milliseconds; ms) a natural logarithm transformation was applied to the three CONVIRT measures. There was no missing data and the assumptions for parametric testing for repeated measures ANOVA, and intra-class correlation were satisfied (TaBrAchnick and Fidell, 2013).

### Test-retest reliability

Test-retest reliability of the CONVIRT measures across the three testing phases were analysed via an intra-class correlation. Cronbach's alpha for SAC-VR, DET-VR, IDN-VR and the COMP were classified as 'excellent' (Cicchetti, 1994; see Table 2). Participants' rankings remained highly stable across the differing BRAC, indicating that the consistent application

**Table 2.** Intra-class correlations of the CONVIRT measures

	SAC-VR	DET-VR	IDN-VR	COMP
ICC <i>r</i> (95CI)	.80 (.70–.88)	.83 (.73–.90)	.83 (.69–.88)	.91(.86–.94)

*Note.* ICC = Intra-Class Correlation Coefficient; CI = Confidence Interval; SAC-VR = Saccade Reaction Time; DET-VR = Choice Reaction Time; IDN-VR = Simple Reaction Time; COMP = Composite Score, an amalgamation of the three CONVIRT measures

of alcohol induced cognitive impairment, fatigue and practice effects uniformly. This suggests that the CONVIRT measures are reliable measures of cognitive performance when cognitive functioning is impaired.

### Assessing change across trials of increasing BRAC

Descriptive statistics for the entire sample were calculated from the CONVIRT measures across differing BrAC levels and indicated that the speed of participants' responses appeared to slow as BrAC levels increased (see Table 3).

A series of one-way repeated measures ANOVAs were conducted to assess if each of the CONVIRT measures were sensitive measures of cognitive impairment across BrAC levels. Mauchly's test of sphericity indicated that the assumption of sphericity had not been violated for COMP,  $\chi^2(2) = 0.50$ ,  $p = .780$ . However, sphericity had been violated for SAC-VR,  $\chi^2(2) = 6.10$ ,  $p = .047$ , DET-VR,  $\chi^2(2) = 8.00$ ,  $p = .018$ , and IDN-VR,  $\chi^2(2) = 18.25$ ,  $p < .001$ . Therefore, a Greenhouse–Geisser correction was used for SAC-VR, DET-VR and IDN-VR (TaBrACHnick & Fidell, 2013).

Analyses indicated that there was a deleterious effect of increasing BRAC level on SAC-VR, ( $F(1.80, 95.44) = 4.02$ ,  $p = .025$ ,  $\eta_p^2 = .07$ ), DET-VR, ( $F(1.74, 92.78) = 30.02$ ,  $p < .001$ ,  $\eta_p^2 = .36$ ), IDN-VR, ( $F(1.54, 81.80) = 7.61$ ,  $p = .008$ ,  $\eta_p^2 = .13$ ), and, COMP, ( $F(2, 106) = 26.87$ ,  $p < .001$ ,  $\eta_p^2 = .34$ ). The effect sizes for SAC-VR and IDN-VR were considered moderate, while the DET-VR and COMP effects were large (Cohen, 1988).

To assess which of the comparisons between BrAC phases differed, the Holm's multi-stage method (Holm, 1979) was utilized to assess the three comparisons across differing BrAC levels, ordered by descending effect size, against error rates ( $\alpha$ ) for each CONVIRT measure. The highest effect was provided a criterion  $\alpha$  of .05/3 (.02), followed by the next

highest at .05/2 (.03), and .05/1 (.05). There were significant differences in performance across increasing levels of BrAC on SAC-VR, DET-VR, IDN-VR, and COMP (see Table 2). The effect sizes were moderate to large (Cohen, 1988).

The effect of increasing BrAC to 0.08% resulted in substantially diminished performance from baseline on all CONVIRT tests. This is illustrated (Fig. 3) by the significant linear trend observed from the within-subject contrasts for the SAC-VR ( $F(1, 53) = 10.76$ ,  $p = .002$ ,  $\eta_p^2 = .17$ ), DET-VR ( $F(1, 53) = 42.08$ ,  $p < .001$ ,  $\eta_p^2 = .44$ ), IDN-VR ( $F(1, 53) = 10.36$ ,  $p = .002$ ,  $\eta_p^2 = .16$ ), and COMP ( $F(1, 53) = 48.96$ ,  $p < .001$ ,  $\eta_p^2 = .48$ ).

## Discussion

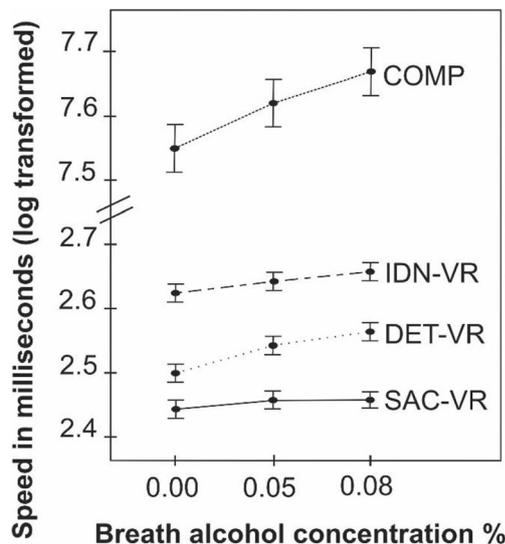
The CONVIRT battery of tests was able to sensitively detect the cognitive impairment induced by alcohol consumption at BrAC% levels at both 0.05 and 0.08 with performance on all tests showing the expected dose–response characteristics. The DET-VR and COMP scores provided the greatest sensitivity to the alcohol challenge as indicated by the large effect sizes detected for the performance measures from these two tests. Additionally, the test–retest reliability of the main performance measures from the CONVIRT battery were by definition, very high (i.e., all  $d > 0.8$ ) and this obviously contributed to the sensitivity of the tests.

**Sensitivity and reliability** Our findings concur with results from similar within-group designs that used alcohol to impair cognition. Specifically, healthy young adults using computer-based assessments (Cogstate) post-alcohol consumption show significant decrements in attention (Maruff et al., 2005; Faletti et al., 2003) and decision-making (Faletti et al., 2003). Additionally, the findings from the present study reveal

**Table 3.** Descriptive statistics and comparisons for each CONVIRT measure across differing BrAC ( $N = 54$ )

Measure	BrAC% phase			Comparisons across BrAC% phases		
	0.00 <sup>T1</sup> <i>M (SD)</i>	0.05 <sup>T2</sup> <i>M (SD)</i>	0.08 <sup>T3</sup> <i>M (SD)</i>	T1 - T2	T1 -T3	T2 - T3
SAC-VR	275 (48)	291 (60)	293 (41)	$p = .047$ $d = 0.27$	$p = .002^*$ $d = 0.30$	$p = .589$ $d = 0.01$
DET-VR	312 (37)	351 (56)	370 (76)	$p < .001^*$ $d = 0.96$	$p < .001^*$ $d = 1.27$	$p = .009^*$ $d = 0.41$
IDN-VR	427 (47)	438 (52)	459 (82)	$p = .032$ $d = 0.26$	$p = .018^*$ $d = 0.39$	$p = .028^*$ $d = 0.24$
COMP	7.56 (.12)	7.64 (.15)	7.68 (.16)	$p < .001^*$ $d = 0.85$	$p < .001^*$ $d = 1.20$	$p = .011^*$ $d = 0.35$

*Note.* SAC-VR=Saccade Reaction Time; DET-VR= Simple Reaction Time; IDN-VR= Choice Reaction Time; COMP = Composite Score; All mean values for the SAC-VR, DET-VR, IDN-VR are log transformed in milliseconds however log transformed variables were used in all parametric analyses, and to calculate the COMP variable, M = Mean; SD = Standard Deviation;  $d$  = Cohen's  $d$ ; \* Mean difference is significant at the alpha adjusted rate using the Holms method



**Fig. 3** Change in within-group mean performance across trials of increasing breath alcohol concentration on the CONVIRT battery of tests. Error bars represent the 95% confidence interval

participants had slower saccadic reaction time speeds in response to alcohol administration and this is consistent with other research (e.g., Roche & King, 2010; Vassallo & Abel, 2002) who reported larger (i.e.,  $d > .80$ ) effects. To our knowledge, this is the first such study to determine the link between alcohol related CNS disruption as indicated by reduced visual processing speed using measures within a VR environment. Given that participants showed significant reductions in performance on the CONVIRT measures at the 0.08% BrAC, and the effects of concussion (2-3 days post-concussion) are said to resemble the cognitive decrements at a BrAC of .10% (Maruff et al., 2001), the CONVIRT measure may prove sensitive in assessing the cognitive symptoms related to concussion.

Oculomotor function may be vulnerable and/or compromised due to cognitive impairment (Taghdiri, Varriano, & Tartaglia, 2017). Given that optic movement is an immediate response that acts as the first line of information in reaching a decision in measures of visual attention, working memory and decision-making; ocular movement including measures of saccade speed, plays an important role on the performance on such tests (Taghdiri et al., 2017). As such, eye-tracking methods for the assessment of mild traumatic brain injury including concussion have recently received research attention (Cifu et al., 2015; Heitger et al., 2009; Maruta et al., 2010; Reppert, Lempert, Glimcher, & Shadmehr, 2015). The findings of reductions in SAC-VR in response to alcohol administration in the present study further contribute to this growing evidence base of the efficacy of such measures in detecting cognitive impairment.

In the CONVIRT battery, the DET-VR test (attention) appeared to be the individual measure most sensitive to the cognitive decrements induced by alcohol consumption, while the

SAC-VR measure was the least sensitive. Whether these differences can be explained by the underlying mechanisms used to perform goal-directed or stimulus-directed attention will require further consideration, and possibly require an assessment of the concurrent impact of changes in the activity of the chemical messenger gamma-aminobutyric acid (GABA) with the cognitive measures (Dry, et al., 2012). That said, the COMP measure was also a more sensitive indicator of cognitive deterioration due to alcohol ingestion than the SAC-VR or IDN-VR measures. A potential explanation for this enhanced sensitivity stems from the fact that the COMP score is an amalgamation of the performance across all subtests, and essentially a product of three times as many observations as any of the individual tests. The COMP score, as a consideration of 'global cognitive function' rather than performance in specific cognitive subdomains, may prove to be the most useful measure of cognitive performance. When this information is coupled with the excellent test-retest reliability of the measures across the three trials, this augers well for the use of the CONVIRT measures in studies of concussion.

The intra-class correlation of each test in the CONVIRT battery was excellent when assessed across the three phases. This suggests that the ranking of each participant was held relatively constant across increasing levels of BrAC, and that the participants were very likely cognitively impaired in the same way and to the same level. Therefore, it can be assumed that this protocol, which was similar to Maruff et al.' (2005) testing of simple reaction time from the Cogstate battery, reliably induced acute cognitive impairment in order to simulate the symptoms of concussion. However, identifying if the CONVIRT tests are sensitive to changes in medically diagnosed concussion status is an important next step in the ongoing testing of the CONVIRT battery.

**Strengths and limitations** In this study, important confounds were controlled statistically or via participant exclusion. Specifically, participants did not have a history of concussion or other injuries or illness that may impact performance, and age was entered as a covariate despite the relatively homogenous age range. However, as the findings emanate from a sample of healthy young adults, more testing is required to ensure the findings hold when assessing jockeys. Similarly, although CONVIRT was designed to minimise vestibular disturbance (Horan et al., 2020) and none of the participants reported any nausea or asked for the testing to be stopped at any point in this and earlier studies, future studies will include a quantitative measure of motion sickness to provide a more robust assessment of how participants respond to the VR environment. Finally, it is important to note that the SAC-VR times are more than just a measure of saccadic latency, as they include a portion of acceleration towards the target. Failure to account for this will result in an overestimation of the effect of alcohol on saccadic latency.

The CONVIRT battery appears to be a highly reliable measure of cognitive impairment and sensitive enough to detect changes in cognitive performance across increasing BrAC levels. However, due to differences in gender, BMI, and ethnicity, varying alcohol dosages were administered to participants and this led to differences in experiment duration between participants (Table 1). The experimental design employed meant that there was a chance that practice and fatigue effects could affect performance. However, the SAC-VR test, with a mixture of exogenous and endogenous processes, minimises the impact of practice effects. Further, others have suggested that practice effects in similar computer-based assessments of simple and choice-reaction time (i.e., Cogstate) are confined to re-testing at 10 min (Falleti, Maruff, Collie, & Darby, 2006). Nevertheless, the reliability of the CONVIRT battery needs to be established at longer test–retest intervals.

Future research should seek to remedy the limitations of the present research by assessing alcohol dose–response patterns in conjunction with BrAC levels and are encouraged to use a no-alcohol placebo group to account for expectancy effects and to delineate the effects of practice/fatigue effects from alcohol effects. Alcohol appears to slow cognitive responses rather than qualitatively change the patterns of response (Vorstius, Radach, & Lang, 2012), but whether the same applies to persons post-concussion is unknown.

## Conclusions

From a health and safety perspective, using state-of-the-art technology to assess return-to-normal CNS function and having participants complete their cognitive assessments in environments that better simulate the demands of their work environment is crucial. The CONVIRT battery has high convergent validity with, and higher ecological validity than, standard neuropsychological measures, with no practice effect (Horan et al., 2020). The present study builds on these findings and suggests that CONVIRT is a reliable and sensitive measure that can detect alcohol dose-dependent changes in acute cognitive impairment. The next step will be using the current findings to quantify the effect of concussion on cognition using this battery.

**Funding** There was no external funding of this research.

The data for this research are available at <https://osf.io/4r8wj/>

## References

- Blekher, T., Beard, J. D., O'Connor, S., Orr, W. E., Ramchandani, V. A., Miller, K., ... Li, T. K. (2002). Response of saccadic eye movements to alcohol in African American and non-Hispanic white college students. *Alcoholism: Clinical and Experimental Research*, 26(2), 232–238.
- Cash, C., Peacock, A., Barrington, H., Sinnett, N., & Bruno, R. (2015). Detecting impairment: sensitive cognitive measures of dose-related acute alcohol intoxication. *Journal of Psychopharmacology*, 29(4), 436–446. <https://doi.org/10.1177/0269881115570080>
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6(4), 284–290. <https://doi.org/10.1037/1040-3590.6.4.284>
- Cifu, D. X., Wares, J. R., Hoke, K. W., Wetzel, P. A., Gitchel, G., & Carne, W. (2015). Differential eye movements in mild traumatic brain injury versus normal controls. *Journal of Head Trauma Rehabilitation*, 30(1), 21–28. <https://doi.org/10.1097/HTR.000000000000036>
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.): Hillsdale, N.J.: L. Erlbaum Associates
- Dry, M. J., Burns, N. R., Nettelbeck, T., Farquharson, A. L., & White, J. M. (2012). Dose-related effects of alcohol on cognitive functioning. *PLoS one*, 7(11), e50977–e50977. <https://doi.org/10.1371/journal.pone.0050977>
- Falleti, M. G., Maruff, P., Collie, A., & Darby, D. G. (2006). Practice effects associated with the repeated assessment of cognitive function using the CogState battery at 10-minute, one week and one month test-retest intervals. *Journal of Clinical and Experimental Neuropsychology*, 28(7), 1095–1112
- Falleti, M. G., Maruff, P., Collie, A., Darby, D. G., & McStephen, M. (2003). Qualitative similarities in cognitive impairment associated with 24h of sustained wakefulness and a blood alcohol concentration of 0.05%. *Journal of Sleep Research*, 12(4), 265–274. <https://doi.org/10.1111/j.1365-2869.2003.00363.x>
- Heitger, M. H., Jones, R. D., Macleod, A. D., Snell, D. L., Frampton, C. M., & Anderson, T. J. (2009). Impaired eye movements in post-concussion syndrome indicate suboptimal brain function beyond the influence of depression, malingering or intellectual ability. *Brain*, 132(Pt 10), 2850–2870. <https://doi.org/10.1093/brain/awp181>
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2), 65–70.
- Horan, B., Heckenberg, R.A., Maruff, P., & Wright, B.J. (2020). Development of a new virtual reality test of cognition: Assessing the test–retest reliability, convergent and ecological validity of CONVIRT. *BMC Psychology*, 8(61), <https://doi.org/10.1186/s40359-020-00429-x>
- Howett, D., Castegnaro, A., Krzywicka, K., Hagman, J., Marchment, D., Henson, R., ... Chan, D. (2019). Differentiation of mild cognitive impairment using an entorhinal cortex-based test of virtual reality navigation. *Brain*, 142(6), 1751–1766.
- Jongen, S., Vuurman, E. F. P. M., Ramaekers, J. G., & Vermeeren, A. (2016). The sensitivity of laboratory tests assessing driving related skills to dose-related impairment of alcohol: A literature review. *Accident Analysis & Prevention*, 89, 31–4.
- King, A. C., & Byars, J. A. (2004). Alcohol-induced performance impairment in heavy episodic and light social drinkers. *Journal of Studies on Alcohol and Drugs*, 65(1), 27–36. <https://doi.org/10.15288/jsa.2004.65.27>
- Maruff, P., Collie, A., Darby, D.G. (2001). Quantifying the cognitive impairment associated with concussion: Using blood alcohol concentration as a reference point. *British Journal of Sports Medicine*, 35(5), 373
- Maruff, P., Falleti, M. G., Collie, A., Darby, D., & McStephen, M. (2005). Fatigue-related impairment in the speed, accuracy and variability of psychomotor performance: comparison with blood alcohol levels. *Journal of Sleep Research*, 14(1), 21–27. <https://doi.org/10.1111/j.1365-2869.2004.00438.x>

- Maruta, J., Suh, M., Niogi, S. N., Mukherjee, P., & Ghajar, J. (2010). Visual tracking synchronization as a metric for concussion screening. *Journal of Head Trauma Rehabilitation, 25*(4), 293–305. <https://doi.org/10.1097/HTR.0b013e3181e67936>
- Mullen, S. J., Yücel, Y. H., Cusimano, M., Schweizer, T. A., Oentoro, A., & Gupta, N. (2014). Saccadic eye movements in mild traumatic brain injury: A pilot study. *Canadian Journal of Neurological Sciences, 41*(1), 58–65.
- Orban de Xivry, J. J., & Lefevre, P. (2007). Saccades and pursuit: two outcomes of a single sensorimotor process. *The Journal of Physiology, 584*(1), 11–23
- Ramat, S., Leigh, R. J., Zee, D. S., & Optican, L. M. (2007). What clinical disorders tell us about the neural control of saccadic eye movements. *Brain, 130*(1), 10–35.
- Reppert, T. R., Lempert, K. M., Glimcher, P. W., & Shadmehr, R. (2015). Modulation of saccade vigor during value-based decision-making. *Journal of Neuroscience, 35*(46), 15369–15378. <https://doi.org/10.1523/JNEUROSCI.2621-15.2015>
- Roche, D. J., & King, A. C. (2010). Alcohol impairment of saccadic and smooth pursuit eye movements: impact of risk factors for alcohol dependence. *Journal of Psychopharmacology, 212*(1), 33–44.
- Salomoni, P., Prandi, C., Roccetti, M., Casanova, L., Marchetti, L., & Marfia, G. (2017). Diegetic user interfaces for virtual environments with HMDs: A user experience study with oculus rift. *Journal on Multimodal User Interfaces, 11*(2), 173–184. <https://doi.org/10.1007/s12193-016-0236-5>
- Saunders, J. B., Aasland, O. G., Babor, T. F., de la Fuente, J. R., & Grant, M. (1993). Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption—II. *Addiction, 88*(6), 791–804. <https://doi.org/10.1111/j.1360-0443.1993.tb02093.x>
- Schechtman, E., & Shinar, D. (2011). An analysis of alcohol breath tests results with portable and desktop breath testers as surrogates of blood alcohol levels. *Accident Analysis & Prevention, 43*(6), 2188–2194.
- Selin, K. H. (2003). Test–retest reliability of the alcohol use disorder identification test in a general population sample. *Alcoholism: Clinical and Experimental Research, 27*(9), 1428–1435. <https://doi.org/10.1097/01.Alc.0000085633.23230.4a>
- Shields, A. L., & Caruso, J. C. (2003). Reliability generalization of the Alcohol Use Disorders Identification Test. *Educational and Psychological Measurement, 63*(3), 404–413. <https://doi.org/10.1177/0013164403063003004>
- Smith, T. (2011). Breath analysis: Clinical research to the end-user market. *Journal of Breath Research, 5*(3), 032001. <https://doi.org/10.1088/1752-7155/5/3/032001>
- Snegireva, N., Derman, W., Patricios, J., & Welman, K. E. (2018). Eye tracking technology in sports-related concussion: a systematic review and meta-analysis. *Physiological Measurement, 39*(12), 12TR01.
- Sorbello, J. G., Devilly, G. J., Allen, C., Hughes, L. R. J., & Brown, K. (2018). Fuel-cell breathalyser use for field research on alcohol intoxication: an independent psychometric evaluation. *PeerJ, 6*, e4418. <https://doi.org/10.7717/peerj.4418>
- TaBrAChnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics* (6th ed.): Harlow, United Kingdom: Pearson Education Limited.
- Taghdiri, F., Varriano, B., & Tartaglia, M. C. (2017). Assessment of oculomotor function in patients with postconcussion syndrome: A systematic review. *Journal of Head Trauma Rehabilitation, 32*(5), E55–E67. <https://doi.org/10.1097/HTR.0000000000000286>
- Turner, M., McCrory, P., & Halley, W. (2002). Injuries in professional horse racing in Great Britain and the Republic of Ireland during 1992–2000. *British Journal of Sports Medicine, 36*(6), 403–409. <https://doi.org/10.1136/bjism.36.6.403>
- Vassallo, S., & Abel, L. A. (2002). Ethanol effects on volitional versus reflexive saccades. *Clinical & Experimental Ophthalmology, 30*(3), 208–212.
- Vorstius, C., Radach, R., & Lang, A. R. (2012). Effects of acute alcohol intoxication on automated processing: evidence from the double-step paradigm. *Journal of Psychopharmacology, 26*(2), 262–272.
- Vorstius, C., Radach, R., Lang, A. R., & Riccardi, C. J. (2008). Specific visuomotor deficits due to alcohol intoxication: Evidence from the pro- and antisaccade paradigms. *Journal of Psychopharmacology, 196*(2), 201–210. <https://doi.org/10.1007/s00213-007-0954-1>
- Wegner, A. J., & Fahle, M. (1999). Alcohol and visually guided saccades: gap effect and predictability of target location. *Journal of Psychopharmacology, 146*(1), 24–32. <https://doi.org/10.1007/s002130051084>
- Wiener, J. M., Carroll, D., Moeller, S., Bibi, I., Ivanova, D., Allen, P., & Wolbers, T. (2019). A novel virtual-reality-based route-learning test suite: Assessing the effects of cognitive aging on navigation. *Behavior Research Methods, 52*, 1–11.
- Zuba, D. (2008). Accuracy and reliability of breath alcohol testing by handheld electrochemical analysers. *Forensic Science International, 178*, e29–33. <https://doi.org/10.1016/j.forsciint.2008.03.002>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.