ADDICTION
SSA SOCIETY FOR THE STUDY OF ADDICTION
METHODS AND TECHNIQUES
doi:10.1111/add.15563

Raising the bar: improving methodological rigour in cognitive alcohol research

Charlotte R. Pennington[1], Andrew Jones[2], James E. Bartlett[3], Amber Copeland[4] & Daniel J. Shaw[1]

School of Psychology, Aston University, Birmingham, UK,[1] Institute of Population Health Sciences, University of Liverpool, Liverpool, UK,[2] School of Psychology, Arden University, Coventry, UK[3] and Department of Psychology, University of Sheffield, Sheffield, UK[4]

ABSTRACT

Background and Aims A range of experimental paradigms claim to measure the cognitive processes underpinning alcohol use, suggesting that heightened attentional bias, greater approach tendencies and reduced cue-specific inhibitory control are important drivers of consumption. This paper identifies methodological shortcomings within this broad domain of research and exemplifies them in studies focused specifically on alcohol-related attentional bias. **Argument and analysis** We highlight five main methodological issues: (i) the use of inappropriately matched control stimuli; (ii) opacity of stimulus selection and validation procedures; (iii) a credence in noisy measures; (iv) a reliance on unreliable tasks; and (v) variability in design and analysis. This is evidenced through a review of alcohol-related attentional bias (64 empirical articles, 68 tasks), which reveals the following: only 53% of tasks use appropriately matched control stimuli; as few as 38% report their stimulus selection and 19% their validation procedures; less than 28% used indices capable of disambiguating attentional processes; 22% assess reliability; and under 2% of studies were pre-registered. **Conclusions** Well-matched and validated experimental stimuli, the development of reliable cognitive tasks and explicit assessment of their psychometric properties, and careful consideration of behavioural indices and their analysis will improve the methodological rigour of cognitive alcohol research. Open science principles can facilitate replication and reproducibility in alcohol research.

Keywords Addiction, alcohol, attentional bias, cognition, methodology, open science, reliability.

Correspondence to: Charlotte R. Pennington, School of Psychology, College of Health and Life Sciences, Aston University, Birmingham B4 7ET, UK.
E-mail: c.pennington@aston.ac.uk
Submitted 22 October 2020; initial review completed 29 January 2021; final version accepted 28 April 2021

ALCOHOL-RELATED COGNITIONS AND THEIR IMPORTANCE

Dual-process models of addiction propose that the loss of control over alcohol consumption results from an imbalance between two competing systems: an automatic 'impulsive' system triggered by substance-related cues and a more controlled 'reflective' system underpinned by executive functioning [1–3]. According to these models, alcohol misuse develops when the impulsive system becomes hypersensitive through repeated exposure to the rewarding effects of alcohol, which compromises self-control and leads to dysregulated approach bias towards alcohol-related cues [4].

There is considerable interest in identifying and measuring the cognitive processes that drive alcohol (mis)use, not least because this might tell us how alcohol use disorders develop and persist. A wealth of research suggests that heightened attentional bias, greater approach tendencies and reduced cue-specific inhibitory control are important drivers of alcohol consumption and related behaviours (e.g. substance-seeking [4–6]). These distinct but inter-related processes have been shown to predict progression from heavy drinking to dependency [7–9] and the likelihood of relapse following treatment ([10,11]; but see [12]). Moreover, rather than representing stable traits, they appear to fluctuate in response to internal and environmental demands [13,14]. At first glance, these findings have clear health implications; interventions that target these fluctuations effectively might mitigate alcohol-related harm [15].

Importantly, however, several methodological shortcomings cast doubt over the robustness of findings from cognitive alcohol research. In this Methods and techniques article we draw attention to five main issues: (i) a frequent use of inappropriately matched control stimuli; (ii) the

opacity of stimulus selection and validation procedures; (iii) a credence in noisy measures; (iv) a reliance on unreliable tasks; and (v) considerable variability in design and analysis. To exemplify this, we systematically review the last 10 years of literature on one specific subdomain of cognitive alcohol research; namely, alcohol-related attentional bias (n = 64 articles, 68 tasks; https://osf.io/x7gcq/). As shown in Figure 1, this revealed that these issues were present in the majority of synthesized studies. After discussing their respective impact, we then present an easy-to-implement practical guide with a view to establishing gold standards for future research. It is important to stress that many of the issues highlighted in the sections below are applicable beyond the field of cognitive alcohol research and have been discussed within psychological science more generally (e.g. [16–20]). However, it is important to look critically at our own specific field(s) to highlight particular areas in need of methodological reform, and to promote best practices going forward.

METHODOLOGICAL ISSUES

Use of inappropriately matched control stimuli

To investigate alcohol-related cognitions, researchers typically employ experimental paradigms that contrast responses to two categories of stimuli: alcohol-related versus -unrelated. In our review of alcohol-related attentional bias, for example, 61.76% employed the addiction Stroop task [21] or visual probe task (VPT [22]). The former is an adaptation of the emotional Stroop task [23], whereby individuals are required to identify the colour of words that are semantically related or unrelated to alcohol. When colour identification is slower for alcohol-related compared to -unrelated words, this is interpreted as heightened attentional capture by alcohol-related cues [24–26]. During the VPT, individuals are required to respond to a neutral cue (probe) that appears in a location occupied previously by alcohol-related or -unrelated pictorial stimuli. Faster responses to probes appearing in the same location as the former stimulus category are interpreted as attentional bias towards alcohol [27–29]. Both tasks appear to demonstrate construct validity; they generate indices of attentional bias that are associated with individual differences in self-reported alcohol consumption and transient changes in drinking motivation [30–32].

In order to claim that preferential responses towards alcoholic cues on these tasks reflect alcohol-specific attentional bias, it is necessary to employ appropriately matched control stimuli—that is, a category of non-alcoholic substances with some degree of incentive value (e.g. soft drinks). However, our focused review of alcohol-related attentional bias revealed that, of those papers reporting explicitly the stimuli employed within tasks, 35.29% used unmatched non-appetitive control stimuli. For instance, responses towards pictorial alcohol-related stimuli were compared frequently against those to household objects or office stationery (e.g. [33–35]), thereby confounding the incentive value of alcoholic and non-alcoholic appetitive substances. Similarly, although researchers take care with certain validation procedures in the addiction Stroop task, such as ensuring that word length, syllables and frequency of use are matched between experimental and control stimuli, many compare responses between alcohol-related (e.g. VODKA) and non-appetitive words (e.g. CHAIR [36–38]). One study even reports that they employed office stationery purposefully 'so that participants would not be distracted by the control category stimuli in any way' ([39], p. 2). While these studies have

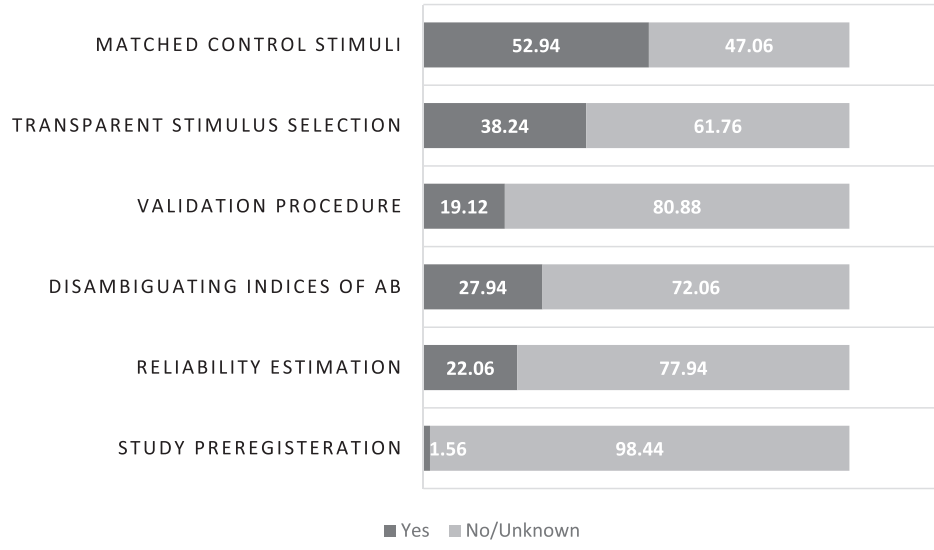

Figure 1 Percentage of tasks (total n = 68) that met the proposed methodological standards. AB = attentional bias

provided important contributions by demonstrating attentional bias towards alcohol-related relative to non-appetitive stimuli, differential responding between these stimulus categories might simply reflect a general (alcohol-unspecific) bias to appetitive stimuli [40,41]. Comparisons among such stimuli that differ markedly in terms of their incentive value therefore make it impossible to isolate the precise mechanisms driving alcohol (mis-) use and may inflate effect size estimates [41–43].

Opacity of stimuli selection and validation

A related but separate issue that casts doubt over the robustness of findings in alcohol research generally is the failure of many studies to report the selection and validation of experimental stimuli. Our review of alcohol-related attentional bias revealed that 11.76% of articles do not describe control stimuli with sufficient detail to evaluate their appropriateness, instead using ambiguous terms such as 'neutral' stimuli. Of the articles that report such information, only 38.24% disclose the source from which their stimuli were selected and only 19.12% report a validation procedure. Moreover, despite the availability of validated image databases, such as the Amsterdam Beverage Picture Set [44], our review indicates that these were utilized by only 15.38% of the studies reporting their source.

Instead, the majority (61.54%) report using stimuli from previous studies but neglect to detail any validity checks. This creates a 'rabbit-hole' problem for researchers in the many instances where materials are not openly available. As one example, the authors of a study published in 2019 cited their previous 2015 article as the origin of the alcohol stimuli, but that article then cites Hogarth *et al.* [45], who employed smoking cues. Such dead ends stifle progress within this research field; researchers are unable to use the same stimuli in order to build upon prior findings, and direct replications are impossible if researchers are forced to develop their own stimuli. Others report using internet image searches to develop stimulus sets, with no information provided about their visual properties (e.g. luminosity) or, therefore, the equivalence between experimental and control stimuli. It is well known that the visual characteristics of stimuli can influence general cognitive processing [28,46], meaning that it is important to standardize stimulus sets in order to reduce noise from these factors. Just as the lack of transparency constrains progress, the frequent disregard for stimulus validation limits the evidential value of cognitive alcohol research.

A credence in noisy measures

Researchers often rely upon measurement indices from raw behavioural data, such as average reaction times (RT) or choice accuracy. For example, our review reveals that behavioural RT was the primary index for 72.06% of measures of attentional bias (with the remaining 27.94% utilizing eye-tracking methodology that can disambiguate attentional processes). This assumes that systematic differences in RT are driven only by attentional bias, but there is a general understanding that RT measures are affected by several cognitive and motor processes simultaneously [47,48]. Specifically, a participant must first encode the stimulus, process information needed to make a decision and then execute an appropriate motor response (e.g. key press). Measurement noise is exacerbated by the fact that common experimental tasks are often unable to account sufficiently for speed–accuracy trade-offs (SATOs [49,50];); while some people will respond faster at the cost of being less accurate, others will respond more slowly to increase their accuracy [51]. By failing to account for SATOs, inferences drawn from raw behavioural data might lack insight into important aspects of the decision-making process (e.g. response caution [49]).

Another issue is the reliance upon subtraction methods (e.g. difference scores) to index alcohol-related cognitions, and assess their associations with other variables of interest (e.g. subjective craving). Difference scores appear to be a simple and effective method of controlling for general RT and isolating signal in the noise. Unfortunately, however, there is a fundamental shortcoming in the use of difference scores; as the correlation increases between two component measures (e.g. RTs to alcohol-related and -unrelated stimuli), the reliability of their difference score decreases proportionately [49], and potentially meaningful associations with other variables are weakened [17]. Together, then, RTs and difference scores are contaminated by factors extraneous to the cognitive mechanism of interest. As the use of such measures constitutes a norm in this domain, these issues question the extent to which existing research can be viewed as obtaining precise, interpretable and sensitive measures of alcohol-related cognition.

Unreliable tasks equal unreliable inferences

Variability in the stimulus sets used across studies, the number of stimuli and their repetitions and the use of noisy measures of response bias will all impact upon the reliability of experimental tasks and the replicability of research findings. Increasing the number of stimuli is believed generally to increase the internal consistency of a task [52], and a large number of stimuli will help to reduce any habituation effect [53]. This is critical, as the stimuli used in alcohol cognition tasks are assumed to evoke an implicit response (e.g. alcohol-related cues should 'grab' attention). Despite this, some tasks used commonly in alcohol (and addiction) research fail to achieve acceptable levels of internal reliability [54,55]; for example, Ataya *et al.* [54] report

alarmingly low estimates for the VPT (α = 0.00–0.50, mean = 0.18), and although the Stroop task outperformed this in a handful of studies, there was marked variability (α = 0.00–0.98, mean = 0.74). Others have confirmed these findings and suggest that such variability might be attributable to specific task features, namely differences in the stimuli used, procedural flexibility (e.g. randomized versus blocked designs, number of stimuli) and serial versus multiple stimulus presentations [56].

Despite its integral role in effect-size estimates and reproducibility [57], reliability was assessed for only 22.06% of the reviewed attentional bias tasks—13.24% reported internal reliability, 10.29% split-half and 8.82% test-retest (some report a combination). Perhaps most strikingly, of the 68 tasks employed, 47.06% were the VPT (18.75% eye-tracking) and a further 14.71% the addiction Stroop, with only 4.41% reporting acceptable reliability explicitly. The poor psychometric properties of some cognitive measures pose a serious issue for the interpretation of research findings and, again, may hamper scientific progress in this field (see [58]). Without investigating and reporting transparently the reliability of cognitive tasks, it is impossible to delineate whether findings from this field are robust or a result of measurement error [16].

Variability in experimental design and analysis decisions

There is substantial variability and opacity in the measures used to operationalize alcohol-related cognition and many intricate design decisions that affect this further. The addiction Stroop task, for example, can differ in the way it is administered (paper-and-pencil versus computerized), the response type measured (key press versus verbal), the number and type of stimuli presented and the design (block versus mixed). Unless reported transparently, such flexibility in methodological choice is likely to restrict the generalizability of findings across studies. Further, seemingly subtle design modifications can impact upon the psychometric properties of a task [59] and statistical power [60].

Alongside heterogeneous stimuli presentation protocols, there is also evidence to suggest a lack of prescriptive analysis strategies across studies. Jones *et al.* [61] noted considerable flexibility in the way that RT outliers were handled in the VPT. In addition, they demonstrated that analysing the same data using different cut-off values led to different estimates of internal consistency and test-retest reliability. Similarly, Jones *et al.* [62] conducted a systematic review of analysis decisions within alcohol and smoking Stroop studies and estimated that more than 7000 analysis pipelines could be attempted. Although these issues extend to the paradigms and techniques employed in other research domains (e.g. functional magnetic resonance imaging [63]), such flexibility is associated with increased false-positive findings, particularly when paired with selective reporting and publication bias [64,65]. Indeed, Jones *et al.* [62] found that key aspects of design and analysis decisions were not disclosed when employing the addiction Stroop task, and our review of the alcohol-attentional bias literature indicates that only one study (1.56%) reported design and analysis decisions a priori through study preregistration.

RAISING THE BAR: RECOMMENDATIONS FOR RESEARCHERS

We have identified numerous shortcomings in the methods employed commonly within studies of not only alcohol-related attentional bias, but cognitive alcohol research more generally. In pursuit of enhancing methodological rigour in this field, we now propose several easy-to-implement practical recommendations. Figure 2 provides a summary.

First, we recommend the use of appropriately matched and validated experimental stimuli to assess alcohol-related cognitions. Control stimuli must be able to isolate the specific cognitive mechanism(s) under investigation; if the aim is to capture individual differences in alcohol-specific cognitions, we recommend that researchers employ matched alcohol-related and appetitive alcohol-unrelated stimuli (e.g. soft drinks). In situations where attentional bias is believed to be unspecific to alcohol, it might be more suitable to employ both appetitive and non-appetitive control stimuli (see [40]). In either case, there needs to be a clear rationale behind stimulus selection. Furthermore, researchers should report stimulus validation procedures routinely. Where possible, researchers can make use of existing validated stimulus sets [44,66–69], and in situations where this is inappropriate (e.g. cultural differences in drinking preferences, brand familiarity) available guidelines (see [44]) should be utilized to develop new sets.

Secondly, we recapitulate calls for a standard practice of reporting the reliability (and validity) of cognitive tasks [18] within alcohol research. This is essential, given that reliability estimates differ between samples, experimental task parameters and measures. A helpful guide is provided by Parsons *et al.* [18], who suggest that permutation split-half reliability should be estimated for individual trial-level data and test-retest reliability when assessing trait constructs. We encourage a focus upon improving the reliability of certain experimental paradigms, such as the VPT and addiction Stroop task [56,70], with a view to developing consensus guides outlining optimal task parameters (e.g. [67]). It is notable, however, that one study reports failed attempts to improve both the test-retest and internal reliability of the VPT based upon empirical

	Recommendations	Resources
Stimuli	• Matched appropriately (luminosity; valence). • Clear rationale behind selection. • Use existing validated stimulus sets where appropriate. • Use existing instructions to develop and validate new stimulus sets.	(44,66–69) S1 File of (44) provides instructions for developing new stimulus sets.
Task	• Assess factors contributing to reliability. • Develop and refine tasks with acceptable psychometric properties and avoid those with poor reliability. • Use tasks with demonstrated reliability. • Develop consensus guides with recommendations for optimal task parameters.	(18,42,58,61,70–72)
Measures	• Justify choice of measures according to established consensus. • Explicitly acknowledge the limitations of certain indices.	(17,47–50)
Analysis	• Include complementary analysis techniques (e.g., modelling) that can afford greater reliability and offer new insights.	(51,73–80)
Disclosure	• Report stimuli selection, validation, and reliability as standard practice. • Report design and analysis decisions through study preregistration or Registered Reports. • Make materials, scripts, and data openly available.	(16,81–83) www.osf.io

Figure 2 Summary of recommendations for improving the methodological rigour of cognitive alcohol research

recommendations [61]. If a task consistently demonstrates suboptimal psychometric properties, then it should be abandoned in favour of alternative reliable tasks (e.g. visual search and free-viewing tasks [1,7,42,58]). Furthermore, because a cognitive task is reliable does not necessarily mean that it is a valid measure of the construct under investigation; some tasks will be better at providing mechanistic insights into the cognitive processes that drive alcohol (mis)use, and it is these tasks that we should seek to optimize (see [72]).

Thirdly, researchers should explore different ways of analysis which might overcome the limitations inherent in the use of raw RTS and difference scores (see [49] for suggested alternatives). Another option is the application of computational modelling [51,73] to alcohol research. One example is the drift-diffusion model (DDM [74,75]), which performs a principled reconciliation of RT and accuracy data to provide accurate estimates of dissociable cognitive and motor processes (e.g. [76,77]). Empirical research demonstrates that the DDM provides more reliable indices of attentional bias (towards threat) derived from the VPT [78] and new interpretations of previous experimental findings [79]. Interestingly, it has also been shown that researchers can benefit from increased statistical power by applying such decision models to experimental designs, without requiring more trials or participants [76]. Although these techniques are yet to be tested empirically within cognitive alcohol research, a recent theoretical review [80] outlines their potential contribution to this field.

A CASE FOR OPEN SCIENCE

Many of the issues highlighted above are compounded by the non-disclosure of important study characteristics (e.g. stimuli selection, analytical decisions), which threaten replicability and reproducibility. This can be improved simply by the implementation of open science practices. There are several excellent guides to adopting open science [81–83], so we focus herein upon solutions that are relevant particularly to the field of alcohol research.

One solution to the lack of transparency around stimulus sets and task parameters is a move to open materials. Currently, reviews of biomedical and addiction sciences indicate that only 1–3% of articles share their methods or protocols through public repositories [84,85], with higher (yet far from optimal) estimates of 14% in psychology [86]. Moreover, a standard practice of open data (where ethically permissible) will allow findings and inferences to be verified and new models to be applied to advance knowledge. We recommend use of the Open Science Framework (www.osf.io), a platform which permits the storage of materials, experimental scripts and data with a CC BY licence, so that any re-use is attributed to the original author(s). Indeed, it has been shown consistently

that articles adopting open research practices receive more citations and lead to research collaborations [87–89].

Rigour in alcohol research can also be enhanced through study preregistration—a time-stamped proposal that makes transparent all key experimental design and analysis decisions in advance, thereby reducing many researcher degrees of freedom. Preregistration can be initiated for both confirmatory and exploratory research. Current rates of preregistration within the addictive behaviour literature are worryingly low; Adewumi *et al.* [84] report that just 3% of articles in addiction medicine were pre-registered, and our own review of empirical research on alcohol-related attentional bias revealed only one pre-registered study. Despite the benefits, preregistration is not a panacea and requires careful oversight by authors, editors, and reviewers.

Their extension is a Registered Report (RR), whereby authors complete a stage one submission outlining their planned methods and analyses. Should this receive approval through the peer-review process, researchers then receive an In Principle Acceptance (IPA); as long as the authors adhere closely to their protocol, the journal agrees to publish the article regardless of the results. Despite recommendations to implement RRs more widely in this specific research domain [90], at the time of this review only four of the 288 journals offering this publishing format focus primarily upon alcohol and substance-use research. Even in their nascent stage, initial evidence suggests that RRs reduce publication bias [91] and enact higher levels of open data and computational reproducibility [92]. Furthermore, RRs receive more citations than would be expected given the impact factor of the journal in which they are published [93,94].

Overall, the adoption of open science is likely to increase replication and reproducibility in alcohol research (see [95]).

CONCLUSIONS

Methodological shortcomings weaken the robustness of cognitive alcohol research. We provide an easy-to-implement guide to enhance rigour in this field; this includes the use of appropriately matched and validated experimental stimuli, a renewed focus upon the development and refinement of reliable experimental tasks and careful consideration of behavioural indices and their analysis. Moreover, we stress the importance of transparent reporting aided by open science principles: stimulus selection, task reliability and validation procedures should be disclosed as standard practice, and study preregistration, open materials and data should be implemented wherever possible. Establishing these recommendations as a gold standard will facilitate replication and reproducibility, thereby increasing trust in a field that proffers important implications for public health and policy.

Declaration of interests

None.

Author contributions

Charlotte Pennington: Conceptualization; data curation; formal analysis; investigation; methodology; project administration; supervision; validation. **Andrew Jones:** Data curation; formal analysis; investigation; methodology; validation. **James Bartlett:** Investigation. **Amber Copeland:** Investigation. **Daniel Shaw:** Data curation; investigation; methodology. All authors contributed to the original draft, review, and editing.

Data availability statement

All materials associated with this manuscript are publicly available on the Open Science Framework: https://osf.io/x7gcq/

References

1. Di Lemma L. C. G., Field M. Cue avoidance training and inhibitory control training for the reduction of alcohol consumption: a comparison of effectiveness and investigation of their mechanisms of action. *Psychopharmacology* 2017; **234**: 2489–98.
2. Wiers R. W., Stacy A. W. Implicit cognition and addiction: an introduction. *Curr Dir Psychol Sci* 2005; **15**: 1–9.
3. Wiers R. W., Bartholow B. D., van den Wildenberg E., Thush C., Engels R. C. M. E., Sher K. J., *et al.* Automatic and controlled processes and the development of addictive behaviors in adolescents: a review and a model. *Pharmacol Biochem Behav* 2007; **86**: 263–83.
4. Fleming K. A., Bartholow B. D. Alcohol cues, approach bias, and inhibitory control: applying a dual process model of addiction to alcohol sensitivity. *Psychol Addict Behav* 2014; **28**: 85–96.
5. Jones A., Robinson E., Duckworth J., Kersbergen I., Clarke N., Field M. The effects of exposure to appetitive cues on inhibitory control: a meta-analytic investigation. *Appetite* 2018; **128**: 271–82.
6. Field M., Munafò M. R., Franken I. H. A. A meta-analytic investigation of the relationship between attentional bias and subjective craving in substance abuse. *Psychol Bull* 2009; **135**: 589–607.
7. Rubio G., Jiménez M., Rodríguez-Jiménez R., Martínez I., Ávila C., Ferre F., *et al.* The role of behavioral impulsivity in the development of alcohol dependence: a 4-year follow-up study. *Alcohol Clin Exp Res* 2008; **32**: 1681–7.
8. van Hemel-Ruiter M. E., Wiers R. W., Brook F. G., de Jong P. J. Attentional bias and executive control in treatment-seeking substance-dependent adolescents: a cross-sectional and follow-up study. *Drug Alcohol Depend* 2016; **159**: 133–41.
9. Wiers C. E., Stelzel C., Park S. Q., Gawron C. K., Ludwig V. U., Gutwinski S., *et al.* Neural correlates of alcohol-approach bias in alcohol addiction: the spirit is willing but the flesh

is weak for spirits. *Neuropsychopharmacology* 2014; **39**: 688–97.
10. Garland E. L., Franken I. H. A., Howard M. O. Cue-elicited heart rate variability and attentional bias predict alcohol relapse following treatment. *Psychopharmacology* 2012; **222**: 17–26.
11. Rupp C. I., Beck J. K., Heinz A., Kemmler G., Manz S., Tempel K., *et al*. Impulsivity and alcohol dependence treatment completion: is there a neurocognitive risk factor at treatment entry? *Alcohol Clin Exp Res* 2016; **40**: 152–60.
12. Christiansen P., Schoenmakers T. M., Field M. Less than meets the eye: reappraising the clinical relevance of attentional bias in addiction. *Addict Behav* 2015; **44**: 43–50.
13. Field M., Werthmann J., Franken I., Hofmann W. The role of attentional bias in obesity and addiction. *Health Psychol* 2016; **35**: 767–80.
14. Jones A., Tiplady B., Houben K., Nederkoorn C., Field M. Do daily fluctuations in inhibitory control predict alcohol consumption? An ecological momentary assessment study. *Psychopharmacology* 2018; **235**: 1487–96.
15. Field M., Christiansen P., Hardman C. A., Haynes A., Jones A., Reid A., *et al*. Translation of findings from laboratory studies of food and alcohol intake into behavior change interventions: the experimental medicine approach. *Health Psychol* 2020; https://doi.org/10.1037/hea0001022
16. Parsons S., Kruijt A.-W., Fox E. Psychological science needs a standard practice of reporting the reliability of cognitive-behavioral measurements. *Adv Methods Pract Psychol Sci* 2019; **2**: 378–95.
17. von Bastian C. C., Blais C., Brewer G. A., Gyukovics M., Hedge C., Kałamała P. *et al*. Advancing the understanding of individual differences in attentional control: Theoretical, methodological, and analytical considerations. 2020 https://doi.org/10.31234/osf.io/x3b9k
18. Verbruggen F., Aron A. R., Band G. P. H., Beste C., Bissett P. G., Brockett A. T., *et al*. A consensus guide to capturing the ability to inhibit actions and impulsive behaviors in the stop-signal task. *eLife* 2019; **8**: e46323.
19. Munafò M. R., Nosek B. A., Bishop D. V. M., Button K. S., Chambers C. D., Percie Du Sert N., *et al*. A manifesto for reproducible science. *Nat Hum Behav* 2017; **1**: 1–9.
20. Silberzahn R., Uhlmann E. L., Martin D. P., Anselmi P., Aust F., Awtrey E., *et al*. Many analysts, one data set: making transparent how variations in analytic choices affect results. *Adv Methods Pract Psychol Sci* 2018; **1**: 337–56.
21. Cox W. M., Fadardi J. S., Pothos E. M. The addiction-Stroop test: theoretical considerations and procedural recommendations. *Psychol Bull* 2006; **132**: 443–76.
22. MacLeod C., Mathews A., Tata P. Attentional bias in emotional disorders. *J Abnorm Psychol* 1986; **95**: 15–20.
23. Stroop J. R. Studies of interference in serial verbal reactions. *J Exp Psychol* 1935; **18**: 643–62.
24. Field M., Christiansen P., Cole J., Goudie A. Delay discounting and the alcohol Stroop in heavy drinking adolescents. *Addiction* 2007; **102**: 579–86.
25. Flaudias V., Brousse G., de Chazeron I., Planche F., Brun J., Llorca P. M. Treatment in hospital for alcohol-dependent patients decreases attentional bias. *Neuropsychiatr Dis Treat* 2013; **9**: 773–9.
26. Spanakis P., Jones A., Field M., Christiansen P. A Stroop in the hand is worth two on the laptop: superior reliability of a smartphone based alcohol Stroop in the real world. *Subst Use Misuse* 2019; **54**: 692–8.
27. Manchery L., Yarmush D. E., Luehring-Jones P., Erblich J. Attentional bias to alcohol stimuli predicts elevated cue-induced craving in young adult social drinkers. *Addict Behav* 2017; **70**: 14–7.
28. Miller M. A., Fillmore M. T. Persistence of attentional bias toward alcohol-related stimuli in intoxicated social drinkers. *Drug Alcohol Depend* 2011; **117**: 184–9.
29. Ramirez J. J., Monti P. M., Colwill R. M. Brief and extended alcohol–cue–exposure effects on craving and attentional bias. *Exp Clin Psychopharmacol* 2015; **23**: 159–67.
30. Díaz-Batanero C., Domínguez-Salas S., Moraleda E., Fernández-Calderón F., Lozano O. M. Attentional bias toward alcohol stimuli as a predictor of treatment retention in cocaine dependence and alcohol user patients. *Drug Alcohol Depend* 2018; **182**: 40–7.
31. Gladwin T. E., Vink M. Alcohol-related attentional bias variability and conflicting automatic associations. *J Exp Psychopath* 2018; **9**: 1–14.
32. Monem R., Fillmore M. T. Alcohol administration reduces attentional bias to alcohol-related but not food-related cues: evidence for a satiety hypothesis. *Psychol Addict Behav* 2019; **33**: 677–84.
33. Ghiță A., Porras García B., Moreno M., Monras M., Ortega L., Mondon S., *et al*. Attentional bias assessment in patients with alcohol use disorder: an eyetracking study. *Ann Rev Cyber Ther Telemed* 2019; **17**: 83–7.
34. Wilcockson T. D. W., Pothos E. M. Measuring inhibitory processes for alcohol-related attentional biases: introducing a novel attentional bias measure. *Addict Behav* 2015; **44**: 88–93.
35. Kim J., Marciano M. A., Ninham S., Zaso M. J., Park A. Interaction effects between the cumulative genetic score and psychosocial stressor on self-reported drinking urge and implicit attentional bias for alcohol: a human laboratory study. *Alcohol Alcohol* 2019; **54**: 30–7.
36. Luehring-Jones P., Louis C., Dennis-Tiwary T. A., Erblich J. A single session of attentional bias modification reduces alcohol craving and implicit measures of alcohol bias in young adult drinkers. *Alcohol Clin Exp Res* 2017; **41**: 2207–16.
37. Rettie H. C., Hogan L. M., Cox W. M. Negative attentional bias for positive recovery-related words as a predictor of treatment success among individuals with an alcohol use disorder. *Addict Behav* 2018; **84**: 86–91.
38. Snelleman M., Schoenmakers T. M., van de Mheen D. Attentional bias and approach/avoidance tendencies do not predict relapse or time to relapse in alcohol dependency. *Alcohol Clin Exp Res* 2015; **39**: 1734–9.
39. Brown C. E., Wilcockson T. D. W., Lunn J. Does sleep affect alcohol-related attention bias? *J Subst Abuse* 2020; **25**: 515–8.
40. Monk R. L., Qureshi A., Pennington C. R., Hamlin I. Generalised inhibitory impairment to appetitive cues: from alcoholic to non-alcoholic visual stimuli. *Drug Alcohol Depend* 2017; **180**: 26–32.
41. Pennington C. R., Qureshi A. W., Monk R. L., Greenwood K., Heim D. Beer? Over here! Examining attentional bias towards alcoholic and appetitive stimuli in a visual search eye-tracking task. *Psychopharmacology* 2019; **236**: 3465–76.
42. Pennington C. R., Shaw D. J., Adams J., Kavanagh P., Reed H., Robinson M., *et al*. Where's the wine? Heavy social drinkers show attentional bias towards alcohol in a visual conjunction search task. *Addiction* 2020; **115**: 1650–9.

43. Versace F., Engelmann J. M., Deweese M. M., Robinson J. D., Green C. E., Lam C. Y., *et al*. Beyond cue reactivity: non-drug-related motivationally relevant stimuli are necessary to understand reactivity to drug-related cues. *Nicotine Tob Res* 2017; **19**: 663–9.
44. Pronk T., van Deursen D. S., Beraha E. M., Larsen H., Wiers R. W. Validation of the Amsterdam beverage picture set: a controlled picture set for cognitive bias measurement and modification paradigms. *Alcohol Clin Exp Res* 2015; **39**: 2047–55.
45. Hogarth L., Dickinson A., Duka T. Detection versus sustained attention to drug cues have dissociable roles in mediating drug seeking behavior. *Exp Clin Psychopharmacol* 2009; **17**: 21–30.
46. Harrison N. R., McCann A. The effect of colour and size on attentional bias to alcohol-related pictures. *Psicológica* 2014; **35**: 39–48.
47. Hedge C., Powell G., Bompas A., Vivian-Griffiths S., Sumner P. Low and variable correlation between reaction time costs and accuracy costs explained by accumulation models: meta-analysis and simulations. *Psychol Bull* 2018; **144**: 1200–27.
48. Miller J., Ulrich R. Mental chronometry and individual differences: modeling reliabilities and correlations of reaction time means and effect sizes. *Psychon Bull Rev* 2013; **20**: 819–58.
49. Draheim C., Mashburn C. A., Martin J. D., Engle R. W. Reaction time in differential and developmental research: a review and commentary on the problems and alternatives. *Psychol Bull* 2019; **145**: 508–35.
50. Heitz R. P. The speed-accuracy tradeoff: history, physiology, methodology, and behavior. *Front Neurosci* 2014; **8**: 1–19.
51. Dutilh G., Annis J., Brown S. D., Cassey P., Evans N. J., Grasman R. P. P. P., *et al*. The quality of response time data inference: a blinded, collaborative assessment of the validity of cognitive models. *Psychon Bull Rev* 2019; **26**: 1051–69.
52. Hoekstra R., Vugteveen J., Warrens M. J., Kruyen P. M. An empirical analysis of alleged misunderstandings of coefficient alpha. *Intern J Soc Res Meth* 2019; **22**: 351–64.
53. Hall G., Rodríguez G. Habituation and conditioning: salience change in associative learning. *J Exp Psychol Anim Learn Cogn* 2017; **43**: 48–61.
54. Ataya A. F., Adams S., Mullings E., Cooper R. M., Attwood A. S., Munafò M. R. Internal reliability of measures of substance-related cognitive bias. *Drug Alcohol Depend* 2012; **121**: 148–51.
55. Field M., Christiansen P. Commentary on Ataya *et al*. (2012), internal reliability of measures of substance-related cognitive bias. *Drug Alcohol Depend* 2012; **124**: 189–90.
56. Christiansen P., Mansfield R., Duckworth J., Field M., Jones A. Internal reliability of the alcohol-related visual probe task is increased by utilising personalised stimuli and eye-tracking. *Drug Alcohol Depend* 2015; **155**: 170–4.
57. Baugh F. Correcting effect sizes for score reliability. *J Appl Psychol* 2002; **62**: 254–63.
58. Soleymani A., Ivanov Y., Mathot S., de Jong P. J. Free-viewing multi-stimulus eye tracking task to index attention bias for alcohol versus soda cues: satisfactory reliability and criterion validity. *Addict Behav* 2020; **100**: 106117.
59. Cooper S. R., Gonthier C., Barch D. M., Braver T. S. The role of psychometrics in individual differences research in cognition: a case study of the AX-CPT. *Front Psychol* 2017; **8**: 1–16.
60. Baker D. H., Vilidaite G., Lygo F. A., Smith A. K., Flack T. R., Gouws A. D., *et al*. Power contours: optimising sample size and precision in experimental psychology and human neuroscience. *Psychol Methods* 2020; https://doi.org/10.1037/met0000337
61. Jones A., Christiansen P., Field M. Failed attempts to improve the reliability of the alcohol visual probe task following empirical recommendations. *Psychol Addict Behav* 2018; **32**: 922–32.
62. Jones A., Worrall S., Rudin L., Duckworth J. J., Christiansen P. May I have your attention, please? Methodological and analytical flexibility in the addiction stroop. *Addict Res Theory* 2021; 1–14.
63. Botvinik-Nezer R., Holzmeister F., Camerer C. F., Dreber A., Huber J., Johannesson M., *et al*. Variability in the analysis of a single neuroimaging dataset by many teams. *Nature* 2020; **582**: 84–8.
64. Simmons J. P., Nelson L. D., Simonsohn U. False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychol Sci* 2011; **22**: 1359–66.
65. Young N. S., Ioannidis J. P. A., Al-Ubaydli O. Why current publication practices may distort science. *PLOS Med* 2008; **5**: 1418–22.
66. Peterson H., Simpson S. L., Laurienti P. J. Wake Forest alcohol imagery set: development and validation of a large standardized alcohol imagery dataset. *Alcohol Clin Exp Res* 2019; **43**: 2559–67.
67. Onie S., Gong S., Manwaring E., Grageda D., Webb K., Yuen W. S., *et al*. Validation of the Australian beverage picture set: a controlled picture set for cognitive bias measurement and modification paradigms. *Aust J Psychol* 2020; **72**: 223–32.
68. López-Caneda E., Carbia C. The Galician beverage picture set (GBPS): a standardized database of alcohol and non-alcohol images. *Drug Alcohol Depend* 2018; **184**: 42–7.
69. Stauffer C. S., Dobberteen L., Woolley J. D. American alcohol photo stimuli (AAPS): a standardized set of alcohol and matched non-alcohol images. *Am J Drug Alcohol Abuse* 2017; **43**: 647–55.
70. Grafton B., Teng S., MacLeod C. Two probes and better than one: development of a psychometrically reliable variant of the attentional probe task. *Behav Res Ther* 2021; **138**: 103805.
71. Heitmann J., Jonker N. C., Jong P. J. De, Gladwin TE. A promising candidate to reliably index attentional bias toward alcohol cues: an adapted odd-one-out visual search task. *Front Psychol* 2021; **12**: 1–11.
72. Hedge C., Bompas A., Sumner P. Task reliability considerations in computational psychiatry. *Biol Psychiatry Cogn Neurosci Neuroimag* 2020; **5**: 837–9.
73. Guest O., Martin A. E. How computational modeling can force theory building in psychological science. *Perspect Psychol Sci* 2021; https://doi.org/10.1177/1745691620970585
74. Ratcliff R., McKoon G. The diffusion decision model: theory and data for two-choice decision tasks. *Neural Comput* 2008; **20**: 873–922.
75. Ratcliff R., Smith P. L., Brown S. D., McKoon G. Diffusion decision model: current issues and history. *Trends Cogn Sci* 2016; **20**: 260–81.
76. Stafford T., Pirrone A., Croucher M., Krystalli A. Quantifying the benefits of using decision models with response time and accuracy data. *Behav Res Methods* 2020; **52**: 2142–55.
77. Lerche V., Voss A. Retest reliability of the parameters of the Ratcliff diffusion model. *Psychol Res* 2017; **81**: 629–52.

78. Price R. B., Brown V., Siegle G. J. Computational modeling applied to the dot-probe task yields improved reliability and mechanistic insights. *Biol Psychiatry* 2019; **85**: 606–12.
79. Pirrone A., Dickinson A., Gomez R., Stafford T., Milne E. Understanding perceptual judgment in autism spectrum disorder using the drift diffusion model. *Neuropsychology* 2017; **31**: 173–80.
80. Field M., Heather N., Murphy J. G., Stafford T., Tucker J. A., Witkiewitz K. Recovery from addiction: behavioral economics and value-based decision making. *Psychol Addict Behav* 2020; **34**: 182–93.
81. Kathawalla U.-K., Silverstein P., Syed M. Easing into open science: a guide for graduate students and their advisors. *PsyArXiv* 2020; https://doi.org/10.31234/osf.io/vzjdp
82. Soderberg C. K. Using OSF to share data: a step-by-step guide. *Adv Methods Pract Psychol Sci* 2018; **1**: 115–20.
83. Kiyonaga A., Scimeca J. M. Practical considerations for navigating registered reports. *Trends Neurosci* 2019; **42**: 568–72.
84. Adewumi M. T., Vo N., Tritz D., Beaman J., Vassar M. An evaluation of the practice of transparency and reproducibility in addiction medicine literature. *Addict Behav* 2021; **112**: 106560.
85. Iqbal S. A., Wallach J. D., Khoury M. J., Schully S. D., Ioannidis J. P. A. Reproducible research practices and transparency across the biomedical literature. *PLOS Biol* 2016; **14**: 1–13.
86. Hardwicke T., Thibault R., Kosie J., Wallach J., Kidwell M., Ioannidis J. Estimating the prevalence of transparency and reproducibility-related research practices in psychology (2014–2017). *MetaArXiv* 2020; https://doi.org/10.31222/osf.io/9sz2y
87. Allen C., Mehler D. M. A. Open science challenges, benefits and tips in early career and beyond. *PLOS Biol* 2019; **17**: 1–14 (e3000246).
88. Piwowar H. A., Vision T. J. Data reuse and the open data citation advantage. *Peer J* 2013; **2013**: 1–25.
89. McKiernan E. C., Bourne P. E., Brown C. T., Buck S., Kenall A., Lin J., *et al*. How open science helps researchers succeed. *eLife* 2016; **5**: 1–19.
90. Gorman D. M. Use of publication procedures to improve research integrity by addiction journals. *Addiction* 2019; **114**: 1478–86.
91. Scheel A., Schijen M., Lakens D. An excess of positive results: comparing the standard psychology literature with registered reports. *Adv Methods Pract Psychol Sci* 2020; https://doi.org/10.1177/251524592117467
92. Obels P., Lakens D., Coles N. A., Gottfried J., Green S. A. Analysis of open data and computational reproducibility in registered reports in psychology. *Adv Methods Pract Psychol Sci* 2020; **3**: 229–37.
93. Chambers C. The registered reports revolution: lessons in cultural reform. *Significance* 2019; **16**: 23–7.
94. Chambers C., Tzavella L. Registered reports: past, present and future. *MetaArXiv* 2020; https://doi.org/10.31222/osf.io/43298
95. Protzko J., Krosnick J., Nelson L., Nosek B., Axt J., Berent M., *et al*. High replicability of newly-discovered social-behavioral findings is achievable. *PsyArXiv* 2020; https://doi.org/10.31234/osf.io/n2a9x

