ARE THERE SEX DIFFERENCES IN THE RELATIONSHIP BETWEEN HEAVY ALCOHOL USE AND DISINHIBITION? A META-ANALYSIS

Janette L. Smith and Richard P. Mattick

Technical Report Number 339

Funded by an Australian Rotary Health
Postdoctoral Research Fellowship to Dr Janette Smith

ISBN: 0-7334-3822-9

Table of Contents

Executive Summary

Background: Deficits in behavioural inhibitory control are reported in substance abuse, including alcohol, yet few studies consider the possibility of sex differences in the relationship between heavy alcohol use and disinhibition. Here, we meta-analyse several studies to determine whether there are sex differences in the relationship between heavy alcohol use and disinhibition.

Methods: We used random-effects models to integrate results from 31 studies that compared alcohol dependent or heavy drinker groups with healthy control participants, to estimate the effect sizes separately for males and females. We consider results from the Go/NoGo and stop-signal tasks separately, with the most important variables being the proportion of commission errors to NoGo stimuli in the Go/NoGo task, and the stop-signal reaction time (SSRT) in the stop-signal task.

Results: For heavy drinkers, there was no evidence of a sex difference in the relationship between heavy alcohol use and disinhibition (sex difference in effect sizes for NoGo errors: $p = 0.782$; SSRT: $p = 0.179$), and indeed little evidence of any cognitive deficit associated with risky drinking. For alcohol dependence, evidence for cognitive decrements was greater, however, the pattern of results suggested problems with stimulus discrimination rather than inhibition specifically; further, although effect sizes were larger for women than men, there were no significant sex differences observed (NoGo errors: $p = 0.595$; SSRT: $p = 0.248$). The majority of studies on alcohol dependent groups did not examine sex as a factor in their analyses, or even excluded women from the sample altogether.

Conclusions: Under-representation of women among studies of inhibitory dysfunction in alcohol dependence, and concomitant lack of statistical power, means that the question of a sex-sensitive effect cannot yet be definitively answered. Suggestions are made concerning the need to recruit more women into research studies, and record more information about them, with implications concerning treatment discussed.

1. INTRODUCTION

Widespread evidence exists for sex differences in both levels of alcohol consumption and the effects of alcohol consumption. Men show lower rates of abstention, more frequent alcohol use, and higher prevalence of heavy episodic (binge) drinking than women (World Health Organization, 2011). However, evidence also suggests that substance abuse is a more severe disorder for women: given the opportunity, women are as likely to use drugs as males (Van Etten et al., 1999); among users, rates of dependence are similar for men and women for many drugs (Anthony et al., 1994); and women who are drug dependent use more of the drug and become dependent more quickly than men, that is, the "telescoping" effect (Hser et al., 2005, Piazza et al., 1989). Furthermore, brain volume atrophy and cognitive deficits associated with chronic heavy use develop more quickly in women than men (for a review, see Erol and Karpyak, 2015), and for a given amount of alcohol use, alcohol appears to affect brain structure and function in women more than men (Hommer, 2003, Mann et al., 2005, Sharrett-Field et al., 2013). Since the disorder appears more severe among women, and given the narrowing alcohol consumption gap between men and women (White et al., 2015, Slade et al., 2016), more information is required on sex differences in cognitive dysfunction associated with heavy alcohol use. Here, we focus on one aspect of cognitive function known as inhibitory control. As we shall review, there is some evidence for greater inhibitory deficits among heavy drinking women, but few researchers have examined the possibility of sex differences in inhibitory control among alcohol dependent groups.

The ability to delay, withhold, or interrupt an inappropriate behaviour is core to everyday functioning. Withholding or delaying an immediate inappropriate response allows time for other goal-directed high-order processes, such as formulation of alternate plans of actions, including their possible outcomes, and selection of a more appropriate course of action (Barkley, 1997). Deficits in behavioural inhibition are attracting increasing interest as a contributor to the development and maintenance of substance use disorders, particularly in relation to bingeing and relapse. That is, impulsive disinhibited responses predominate whenever immediately salient, drug-related rewards are present (Goldstein and Volkow, 2002, Hester et al., 2010). We recently confirmed a reduction in inhibitory capacity in regular users of several substances, including alcohol dependence and heavy drinkers (Smith et al., 2014). The current body of evidence suggests that, rather than being the result of substance abuse, such problems with inhibitory control precede and predict the development of substance abuse; longitudinal studies have shown that differences in inhibitory control are apparent in late childhood and adolescence, between those who would later develop substance problems and those who would not (Mahmood et al., 2013, Norman et al., 2011, Tarter et al., 2003). Further, many studies have shown that greater problems with inhibitory control are associated with poorer treatment outcomes, including relapse and treatment dropout (Czapla et al., 2016, Goudriaan et al., 2008, Luijten et al., 2016, Petit et al., 2014, Prisciandaro et al., 2013, Rupp et al., 2016, Spechler et al., 2016, Steele et al., 2014), and treatment adjuncts aimed at improving inhibitory control are currently being investigated, such as administration of modafinil (Mereu et al., 2013) or cognitive training (Allom et al., 2016, Jones et al., 2016). It is clear that disinhibition is an important avenue for further investigation as it relates to all phases of the addiction life-cycle, especially since the presence of sex differences in this relationship may suggest different prevention or treatment options for women and men.

Despite the need for research into sex differences, we noted in our earlier study (Smith et al., 2014) that the literature studied mostly male participants; the sex difference in recruitment was apparent for almost all substances, but especially for alcohol dependence (an effect also reported elsewhere: Lind et al., 2017). Among heavy drinkers, some studies have reported that females showed greater inhibitory decrements than males (Kreusch et al., 2013, Nederkoorn et al., 2009, Smith et al., 2016, Townshend and Duka, 2005, Weafer et al., 2015, Weafer and de Wit, 2014), suggesting that reliance on primarily male participants in dependent samples may underestimate the inhibitory deficit associated with heavy alcohol consumption. On the other hand, there are also several reports of no sex differences in inhibitory capacity among heavy drinkers (Czapla et al., 2015, Franken et al., 2017, Rossiter

et al., 2012). Clearly, more work is needed to reconcile these results. Here, we review and meta-analyse published evidence for a sex difference (or lack thereof) in the relationship between heavy alcohol use and disinhibition, focusing on two experimental paradigms commonly used to assess inhibitory control.

In the Go/NoGo task, participants must press a button to one 'Go' stimulus and withhold that response to another 'NoGo' stimulus. The main measures of interest are the rate of commission errors, that is, incorrect responses to NoGo stimuli (aka false alarms); the rate of omission errors, that is, the absence of a response to Go stimuli (aka misses); and reaction time (RT) to Go stimuli. An increased NoGo error rate in combination with no difference in Go errors (and sometimes a shorter Go RT) is the clearest evidence of a reduction in inhibitory capacity; a high rate of Go errors is thought to reflect problems with sustained attention (Trommer et al., 1988), while increases in all three measures are thought to reflect a general cognitive control decrement (Wright et al., 2014).

Go/NoGo task design parameters such as the frequency of Go stimuli and the rate of stimulus presentation have important effects on the inhibitory demands of the task. In the current study, we focus only on forms of the Go/NoGo task that produce strong inhibitory effects. Specifically, research has shown that inhibition is more difficult when NoGo stimuli are less frequent than Go stimuli, as the Go response becomes prepotent (Nieuwenhuis et al., 2003, Wessel, 2018, Bruin and Wijers, 2002). Further, inhibition is more difficult when the interval between the onset of successive stimuli is relatively short (Zamorano et al., 2014). Indeed, Wessel (2018) has shown that the brain electrical activity "signature" of inhibitory control is most strongly apparent when NoGo stimuli are rare and there is a short interval between stimuli. Therefore, here we consider only studies where Go stimuli are at least 51%, and with a mean stimulus onset asynchrony (SOA) of 2000ms or less.

We also consider results from studies using the stop-signal task (Logan and Cowan, 1984). In this paradigm, participants respond with a left or right button press to two different types of stimuli (e.g., leftward and rightward white arrows), and attempt to stop that response if a 'stop-signal' is presented on a small proportion of trials (e.g., the arrow changes to red). The stop-signal is presented randomly and at variable delays after the Go stimulus, such that inhibition is easy when the stop-signal is presented early, and difficult when presented late. The probability of inhibition at different delays allows estimation of the stop-signal reaction time (SSRT), or the time required to stop a response (Logan, 1994), around 200-250ms for healthy control adults (Band et al., 2003). The SSRT and the RT to Go trials (i.e., those on which no stop-signal is presented) are independent (Logan, 1994); the presence of a difference on one measure does not alter the interpretation of the other, unlike the Go/NoGo task. In our earlier meta-analysis, we showed an inhibitory reduction in both alcohol dependent and heavy drinker groups in the stop-signal task, but in the Go/NoGo task, only alcohol dependent groups showed an inhibitory decrement (Smith et al., 2014).

In the current work, we synthesise the available cross-sectional research to estimate the extent of inhibitory and other cognitive control problems in male and female heavy drinkers and alcohol dependent groups relative to non- or light-drinking controls. We also compare the estimated effect sizes to observe whether inhibitory differences associated with heavy drinking are significantly different for males and females.

2. METHODS

2.1 Search strategy

A literature search with no date restrictions was conducted in the PubMed, PsycINFO, Project Cork, DRUG, Medline, Medline in process, Embase and CINAHL electronic databases by an author and a qualified librarian at the National Drug and Alcohol Research Centre, University of New South Wales, in May 2016; an updated literature search was conducted by JS in November 2017 for papers published in the interim. The search terms used were “Go-NoGo”, “Go/NoGo”, “NoGo”, “SSRT”, “stop-signal task”, “stop-signal”, “SST”, “continuous performance task”, “CPT”, “CANTAB”, “Gordon Diagnostic System”, “response inhibition”, “inhibit” with explosion, “disinhibit” with explosion, “neurocognitive function”, “executive function”, “executive dysfunction”, “cognitive control”, “cognition disorders”, “reaction time” and “behavioural control” in combination with “alcohol”, “alcohol drinking”, “alcohol consumption”, “alcohol use disorder”, “alcoholism”, “heavy drinker”, “binge drinking”, “light drinker”, “social drinker” and “non-drinker”, with exclusion of “NoGo-A” (a neural growth inhibiting protein). Search terms involving the anti-saccade task were also included but are not presented here due to a lack of research investigating this task. The titles, abstracts and texts of the retrieved articles were scanned by JS to determine eligibility. The reference lists of articles that met inclusion criteria were also reviewed for further relevant articles.

2.2 Inclusion/exclusion criteria

Included papers were required to: (a) be presented in full text in English; (b) be conducted on human participants; (c) compare an alcohol-dependent or chronic heavy-drinker group to a control (non-drinker or limited use) group (i.e., cross-sectional data were presented); (d) report at least one measure from the following: SSRT, Go RT in the stop-signal task; NoGo commission errors, Go omission errors, Go RT in the Go/NoGo task; (e) for Go/NoGo task studies, use Go probability of at least 51%, and a mean interval between the onset of successive stimuli of 2000ms or less as described in the Introduction; (f) not report data from samples which overlapped with another included paper; when overlap was indicated either in the published paper or via communication with an author, we used the publication with the larger sample for a more precise estimation of effect size; (g) provide sufficient information for calculating effect sizes separately for each sex (either in the published article, or provided by an author on request, as means, standard deviations and sample sizes separately for men and women for each group). A minimum two attempts were made to contact multiple co-authors of studies which did not report sufficient information for calculation of effect sizes independently for each sex. We did not exclude papers which reported on participants of one sex only, as these could still contribute to estimation of the effect size for that sex (we confirmed that similar weighted mean effect sizes were obtained whether these papers were included or excluded). We excluded papers where the stop-signal was presented at only one delay, and where it indicated that participants should change the initiated response to a different response; we also excluded cued-Go/NoGo tasks since there were too few with informative cues (which increase the need for inhibition: Randall and Smith, 2011) for meta-analysis. We also excluded studies focusing solely on the acute effects of alcohol consumption, or on family members of heavy users of alcohol, including studies of in utero exposure, without comparison between heavy and light-drinking groups.

682 unique abstracts were identified in the literature search and screened; 95 full-text articles were reviewed; 41 articles met inclusion criteria (a)-(f), and of those, we obtained sufficient information for calculation of effect sizes to retain 31 papers for the meta-analysis. Six of the papers reported data from one sex only; authors provided the required information on request for 23 papers, and only two papers included data for each sex in the published article.

2.3 Effect size calculations and pooling strategies

Where performance was reported as accuracy or hit rate, values were transformed to error rate before calculation of effect sizes so that for all measures, a positive effect indicates poorer performance in the heavy drinker/dependent group. For the stop-signal task, we

calculated mean effect sizes for the SSRT and for Go reaction time, where available. For the Go/NoGo task, we calculated mean effect sizes for commission errors, omission errors and Go RT, where available. For each measure, we calculated Hedges' g in line with guidelines from Lipsey and Wilson (2001) and performed with the associated webpage https://www.campbellcollaboration.org/effect-size-calculato.html. In order to correct for overestimation of the effect size associated with small sample sizes, we applied Hedges' (1980) correction to each effect size, and calculated inverse variance weights for each study using the corrected effect size.

Effect sizes for all papers were coded by author JS twice on separate occasions, the second without reference to the first, and the results compared for inconsistencies (the majority of which were minor and resulted from simple data recording/entry errors). Within each analysis (e.g., NoGo errors for alcohol dependence), only one effect size was calculated from each study, although most studies reported more than one measure (e.g., also reporting Go errors for alcohol dependence).

The weighted mean effect size (random effects model) was calculated as the sum of weighted adjusted effect sizes, for each of the k effect sizes, divided by the sum of the weights for all k studies in the analysis. We also calculated the standard error of the mean as well as 95% confidence intervals, and tested the null hypothesis that mean effect size = 0 using a z-score transformation, with significance set at $p < 0.05$, two-tailed. Heterogeneity was assessed with the Q statistic (Hedges and Olkin, 1985), compared for significance against a chi-square distribution with k-1 degrees of freedom. If Q is significant, substantial heterogeneity between effect sizes is indicated for that analysis. Further, we also report the I^2 statistic, the percentage of variation across studies that is associated with heterogeneity rather than chance (Higgins et al., 2003). However, no analysis of moderator variables was possible, since no measure had the required minimum 20 effect sizes recommended for sufficient power for moderator analysis (Marín-Martínez and Sánchez-Meca, 1998, Sánchez-Meca and Marín-Martínez, 1998). Further, we intended to assess publication bias using funnel plot techniques and Egger's regression test (Egger et al., 1997), but could not, as for almost all measures, there were insufficient numbers of included papers to reliably assess these plots or statistics (a minimum of 10 studies are needed for sufficient power, more if there is substantial heterogeneity between studies: Sterne et al., 2011). However, given that the majority of studies included here were interested in examination of group differences (e.g., heavy drinker or alcohol-dependent vs. control), rather than sex differences, it seems unlikely that the presence or absence of sex differences contributed to a decision to publish or not. On the other hand, it is possible (and even probable, if the inhibitory deficit is truly less severe in male abusers of alcohol) that studies with male-skewed samples may find smaller effects, and that researchers may have decided not to publish such studies, independent of any hypotheses about sex differences.

Lastly, we compared the weighted mean effect sizes for men and women using a fixed effects model (Borenstein and Higgins, 2013); specifically, a z score was calculated as the difference between weighted mean effect sizes, divided by the square root of the sum of variances associated with each mean effect size, and assessed for significantly different effect size estimates for men and women by comparing this score against the z distribution.

3. RESULTS

3.1 Heavy drinkers

Table 1 presents information on the individual studies examining heavy and light drinkers performing the Go/NoGo task and the stop-signal task. Tables 2 and 3 list the effect sizes and weights calculated for each study for heavy drinkers and controls in each task. Table 4 presents the weighted mean effect sizes and homogeneity analyses, as well as analyses of sex differences in the heavy drinker vs. control group comparison. We adhere to Cohen's (1992) classification of effect sizes as small (0.2), medium (0.5) and large (0.8).

Of 17 papers investigating heavy drinkers, 16 collected data from both men and women, and 6 reported analysing sex as a factor. Of those, 3 report larger differences between heavy drinkers and controls for females compared to males, while 3 report no such effect (see Table 1). Total sample sizes (pooled across studies included in the meta-analysis) were generally slightly larger for females than males (see Table 4). Heterogeneity was detected across studies for NoGo and Go errors for females; we also observed a general absence of inhibitory problems in heavy drinkers (NoGo errors and SSRT), with non-significant and small effects around d = 0.2. The only result that approached statistical significance was a small negative effect for SSRT in males, indicating marginally *better* performance in the heavy drinker group compared to controls. Lastly, there were no significant sex differences in the effect size estimates for any measure, and indeed no consistent pattern of more positive effect sizes for females compared to males, despite our hypotheses.

3.2 Alcohol dependence

Table 5 presents information on the individual studies examining alcohol dependent groups compared to controls. Tables 6 and 7 list the effect sizes and weights calculated for each study for alcohol dependent and control groups in each task, and Table 8 presents the summary results.

In marked contrast to the analysis approach for studies of heavy drinkers, only one of the 14 papers examining inhibitory control in alcohol dependence examined sex as a factor in their analysis of behavioural performance, and reported no significant effects of sex; of the remaining 13, only 8 collected data from both male and female participants, and in almost all of these, sample sizes are larger for men than women (see Table 5; in Table 8, men outnumber women by as much as 4.8 to 1). Weighted mean effect sizes were almost universally positive, indicating poorer performance in the alcohol dependent group compared to controls; statistically significant reductions in inhibitory capacity were observed for SSRT in the stop-signal task, with a small effect size for men (d = 0.236) and a medium effect size for women (d = 0.484), however, in the Go/NoGo task, difficulties were observed for both NoGo and Go errors, but not Go RT. These were apparent for both men and women; although we note that effect sizes were generally larger for women than men, there were no statistically significant sex differences in effect size estimates. In the Go/NoGo task, statistically significant heterogeneity between studies was detected for NoGo errors for females, and for both SSRT and Go RT in the stop-signal task for males.

Table 1. Characteristics of study samples by behavioural paradigm and study, for papers reporting on heavy drinkers vs. controls. Criteria refers to the diagnostic criteria used for entry into control and heavy drinker groups. Readers are referred to the original papers for further details.

Task/Study	Control group criteria and mean consumption	Heavy drinker group criteria and mean consumption	Sample size, males		Sample size, females		Group x sex interaction examined?	Notes
			Control	Heavy drinker	Control	Heavy drinker		
Go/NoGo task								
Campanella et al. (2017)	AUDIT score <8 Mean AUDIT: 3.2 Mean grams/week: 15	AUDIT score 8+ Mean AUDIT: 17 Mean grams/week: 120	7	7	10	12	No	Neutral means used
Czapla et al. (2015)	Binge drinking score (from Alcohol Use Questionnaire) 16 or less Mean grams/week: 36	Binge drinking score (from Alcohol Use Questionnaire) 24+; did not meet criteria for dependence Mean grams/week: 139	8	8	8	8	Yes; no significant effects or interactions involving sex	Geometric shape means used
Franken et al. (2017)	'Light' category score on Quantity-Frequency-Variability index Mean drinks on single occasion: 1.9, 1.6 drinking days/week Mean grams/week: 30	'Excessive' or 'Very Excessive' score on Quantity-Frequency-Variability index Mean drinks on single occasion: 3.9, 3.5 drinking days/week Mean grams/week: 137	22	21	25	24	Yes, no significant effects or interactions involving sex	-
Kreusch et al. (2013)	AUDIT score <8, abstainers excluded Mean AUDIT (females): 3.3 Mean AUDIT (males): 3.4 Mean grams/week (females): 30 Mean grams/week (males): 11	AUDIT score 11+ Mean AUDIT (females): 15.2 Mean AUDIT (males): 17.7 Mean grams/week (females): 206 Mean grams/week (males): 255	12	19	24	16	Yes; main effect for sex not significant; female drinkers showed greater NoGo errors than controls for an alcohol image condition (not analysed here); males did not show this effect	Experiment 1 data, means pooled across logo for neutral condition used
Kreusch et al. (2014)	AUDIT score 6 or less Mean grams/week: 23	AUDIT score 11 or more Mean grams/week: 185	7	7	8	8	No	-
Petit et al. (2012)	No family history of alcoholism; AUDIT score <12; abstainers excluded Mean AUDIT: 6.5 Means grams/week: 58	No family history of alcoholism; AUDIT score 12+ Mean AUDIT: 16.9 Mean grams/week: 209	10	8	8	9	No	Neutral means used
Rossiter et al. (2012)	AUDIT score <8 Mean AUDIT: 3.7 Mean grams/week: 31	AUDIT score 16+ Mean AUDIT: 19.4 Mean grams/week: 184	15	19	40	11	Yes, no significant effects or interactions involving sex	Neutral means used
Smith et al. (2017)	4+ drinks/occasion less than once a month (including never); abstainers not excluded Mean AUDIT: 2.9	4+ drinks/occasion 1+/month Mean AUDIT: 10.6	18	12	17	13	No	Means pooled across Stroop and Repeat conditions
Watson et al. (2016)	No reported binge episodes (males: 5+ units, females: 4+ units in a 2 hour period) in past 6 months Mean AUDIT: 4.8	8+ binge episodes in past 6 months Mean AUDIT: 9.8	10	3	8	10	No	Neutral shape means used
Stop-signal task								
Bednarski et al. (2012)	Below median on total monthly units, by gender Males: mean 6 units/month Females: mean 3 units/month Mean AUDIT: 2.3 Mean grams/week: 13	Above median on total monthly units, by gender Males: mean 35 units/month Females: mean 25 units/month Mean AUDIT: 7.1 Mean grams/week: 96	5	9	16	11	No	
Bø and Landrø	< 6 units/week	6+ units/week	86	64	192	55	No	Pooled across 0 and 1-5

(2017)	Mean grams/week: 31	Mean grams/week: 127						units/week groups to create controls, and across 6-10, 11-15 and 15+ groups to create heavy drinkers
Hu et al. (2016)	Non-drinkers	Social drinkers; mean 6.9 occasions/month, with 2.7 units/occasion Mean AUDIT: 4.9 Mean grams/week: 60	31	35	24	22	No	-
Kareken et al. (2013)	No detailed results: of 13 participants, 5 drank 2-4 occasions/week, 4 drank 2-4 occasions/month, 3 drank once/month, and one abstained. Of those who drank, all but one had <5 units/occasion; the last drank 6+/occasion weekly.	Non-dependent, non-treatment-seeking, mean 5.4 units/occasion, mean 1.3 heavy occasions/week Mean AUDIT: 9.1 Mean grams/week: 195	6	11	7	7	No	Family history negative group used
Nederkoorn et al. (2009)	Males: < 12.5 units/week (median split by sex) Females: < 11.5 units/week Mean AUDIT (females): 7.0 Mean AUDIT (males): 7.1 Mean grams/week (females): 61 Mean grams/week (males): 51	Males: > 12.5 units/week Females: > 11.5 units/week Mean AUDIT (females): 11.9 Mean AUDIT (males): 14.7 Mean grams/week (females): 174 Mean grams/week (males): 307	15	16	15	15	Yes; in men, SSRT equal for heavy drinkers and controls; in women, heavy drinkers had longer SSRT than controls. No sex effect for errors to Go trials.	Session 1 neutral means used
Papachristou et al. (2012)	No psychiatric or substance abuse disorders except tobacco; AUDIT score < 11 Mean grams/week: 74	No psychiatric or substance abuse disorders except tobacco; AUDIT score 11+ Mean grams/week: 237	6	4	23	9	No	-
Smith et al. (2016)	4+ drinks/occasion less than once a month (including never); abstainers not excluded Mean AUDIT (females): 3.2 Mean AUDIT (males): 3.1	4+ drinks/occasion 1+/month Mean AUDIT (females): 10.1 Mean AUDIT (males): 9.0	20	21	17	13	Yes; female heavy drinkers had longer SSRT than female controls; male heavy drinkers and male controls not significantly different	-
van Duijvenbode et al. (2013)	AUDIT score <8, abstinent mean 4 years Mean AUDIT in past: 2.6 Mean grams/week in past: 44	AUDIT score 16+, abstinent mean 4 years Mean AUDIT in past: 22 Mean grams/week in past: 1605	7	8	-	-	N/A	Average IQ groups used

Note: in an effort to synthesise data concerning alcohol use in heavy drinker and control groups, we have included where available mean scores for the Alcohol Use Disorders Identification Test (AUDIT; Babor et al., 2001; a score of 8 or above represents hazardous and harmful drinking, while a score of 20 or above indicates probable alcohol dependence), and measures of alcohol consumption for each group. Further, since there are differences in the amount of ethanol per unit between different countries, and between studies for units of time, we have converted all studies into grams of ethanol consumed per week. For studies which did not specifically state the amount of ethanol per unit, we used the definition for the country in which the study was conducted. For studies reporting consumption per month, we converted this to consumption per week by multiplying by 12 and dividing by 52. For Bø and Landrø, we used the midpoint of the ranges up to 14 drinks/week, and 15 for the 15+ group, as a conservative estimate.

Table 2. Summary statistics and weights for each study included in the meta-analysis of Go/NoGo task performance in heavy drinkers vs. controls.

	NoGo errors					Go errors					Go RT				
	g	SE	CI lower	CI upper	Weight (%)	g	SE	CI lower	CI upper	Weight (%)	g	SE	CI lower	CI upper	Weight (%)
Males															
Campanella et al. (2017)	0.145	0.535	-0.904	1.194	6.9	-0.443	0.541	-1.504	0.617	26.2	-0.166	0.535	-1.216	0.883	7.2
Czapla et al. (2015)	0.456	0.506	-0.537	1.448	7.7						-0.525	0.509	-1.521	0.472	8.0
Franken et al. (2017)	-0.246	0.306	-0.847	0.354	21.0						0.117	0.305	-0.481	0.716	22.2
Kreusch et al. (2013)	0.264	0.370	-0.462	0.990	14.4						-0.287	0.371	-1.013	0.440	15.1
Kreusch et al. (2014)	0.344	0.538	-0.711	1.400	6.8						-0.690	0.550	-1.769	0.388	6.8
Petit et al. (2012)	0.000	0.474	-0.930	0.930	8.8	0.381	0.479	-0.557	1.319	31.4	0.000	0.474	-0.930	0.930	9.2
Rossiter et al. (2012)	0.401	0.349	-0.283	1.084	16.2						-0.562	0.352	-1.252	0.128	16.7
Smith et al. (2017)	0.371	0.376	-0.365	1.108	13.9	0.673	0.383	-0.077	1.423	42.3	-0.131	0.373	-0.862	0.600	14.9
Watson et al. (2016)	0.620	0.669	-0.692	1.932	4.4										
Total	0.198	0.140	-0.077	0.473	100.0	0.289	0.318	-0.334	0.911	100.0	-0.231	0.144	-0.513	0.050	100.0
Females															
Campanella et al. (2017)	0.517	0.435	-0.336	1.370	10.8	-0.494	0.435	-1.345	0.358	33.7	0.177	0.429	-0.664	1.018	10.1
Czapla et al. (2015)	**1.318**	**0.552**	0.237	2.400	8.5						-0.519	0.508	-1.516	0.477	7.3
Franken et al. (2017)	0.047	0.286	-0.513	0.608	14.4						-0.133	0.286	-0.693	0.428	21.5
Kreusch et al. (2013)	0.000	0.323	-0.633	0.633	13.4						-0.354	0.325	-0.991	0.283	17.1
Kreusch et al. (2014)	**1.193**	**0.543**	0.129	2.256	8.7						-0.178	0.501	-1.160	0.804	7.5
Petit et al. (2012)	0.377	0.490	-0.584	1.338	9.6	**1.170**	**0.526**	0.139	2.200	29.4	-0.376	0.490	-1.337	0.585	7.8
Rossiter et al. (2012)	**-0.747**	**0.348**	-1.430	-0.064	12.8						**0.704**	**0.348**	0.023	1.386	15.1
Smith et al. (2017)	**0.767**	**0.382**	0.020	1.515	12.0	0.400	0.372	-0.329	1.130	36.9	0.076	0.369	-0.647	0.798	13.5
Watson et al. (2016)	-0.418	0.479	-1.357	0.522	9.9										
Total	0.270	0.216	-0.153	0.692	100.0	0.325	0.446	-0.548	1.198	100.0	-0.035	0.140	-0.309	0.239	100.0

Note: Bold type indicates a significant result; positive effect sizes indicate poorer performance by the heavy drinkers compared to controls.

Table 3. Summary statistics and weights for each study included in the meta-analysis of stop-signal task performance in heavy drinkers vs. controls.

	SSRT					Go RT				
	g	SE	CI lower	CI upper	Weight (%)	g	SE	CI lower	CI upper	Weight (%)
Males										
Bednarski et al. (2012)	0.016	0.558	-1.077	1.110	3.9	0.686	0.573	-0.436	1.809	4.0
Bø & Landrø (2017)	**-0.404**	**0.167**	-0.731	-0.077	43.2	-0.114	0.165	-0.438	0.210	48.2
Hu et al. (2016)	-0.204	0.247	-0.689	0.280	19.6	0.306	0.248	-0.181	0.792	21.4
Kareken et al. (2013)	0.577	0.517	-0.436	1.591	4.5	-0.087	0.508	-1.082	0.908	5.1
Nederkoorn et al. (2009)	-0.008	0.359	-0.713	0.696	9.3					
Papachristou et al. (2012)	0.213	0.647	-1.055	1.482	2.9	-0.532	0.656	-1.818	0.755	3.1
Smith et al. (2016)	-0.251	0.314	-0.866	0.364	12.2	-0.068	0.313	-0.681	0.545	13.5
van Duijvenbode et al. (2013)	0.165	0.518	-0.851	1.181	4.5	0.425	0.523	-0.601	1.451	4.8
Total	-0.206	0.110	-0.421	0.009	100.0	0.029	0.115	-0.196	0.253	100.0
Females										
Bednarski et al. (2012)	-0.330	0.394	-1.102	0.443	11.5	-0.331	0.394	-1.103	0.442	8.3
Bø & Landrø (2017)	-0.086	0.153	-0.386	0.214	29.1	0.053	0.153	-0.247	0.353	55.1
Hu et al. (2016)	-0.074	0.295	-0.652	0.505	16.7	0.274	0.297	-0.307	0.855	14.7
Kareken et al. (2013)	-0.291	0.537	-1.344	0.762	7.1	-0.290	0.537	-1.343	0.763	4.5
Nederkoorn et al. (2009)	**0.751**	**0.378**	0.011	1.492	12.2					
Papachristou et al. (2012)	-0.286	0.395	-1.060	0.488	11.4	-0.626	0.401	-1.412	0.159	8.0
Smith et al. (2016)	**0.748**	**0.381**	0.001	1.494	12.0	0.035	0.368	-0.687	0.757	9.5
van Duijvenbode et al. (2013)										
Total	0.053	0.158	-0.257	0.363	100.0	-0.018	0.114	-0.241	0.204	100.0

Note: Bold type indicates a significant result; positive effect sizes indicate poorer performance by the heavy drinkers compared to controls.

Table 4. Weighted mean effect size (g) and heterogeneity analysis by task, outcome measure, and sex for heavy drinkers vs. controls, as well as statistical tests for sex differences in the weighted mean effect size estimates.

					Effect size analysis						Heterogeneity analysis					Sex differences analysis	
Task	Variable	Sex	Total n, control	Total n, heavy drinker	k	g	SE	95% CI lower	95% CI Upper	z	p	τ^2	Q	df	p	I^2	z	p
Go/NoGo	NoGo errors*	Male	109	104	9	0.198	0.140	-0.077	0.473	1.414	0.157	0.000	3.602	8	0.891	0	0.277	0.782
		Female	148	111	9	0.270	0.216	-0.153	0.692	1.251	0.211	0.242	**20.028**	**8**	**0.010**	60		
	Go errors	Male	35	27	3	0.289	0.318	-0.334	0.911	0.908	0.364	0.092	2.858	2	0.240	30	0.066	0.947
		Female	35	34	3	0.325	0.446	-0.548	1.198	0.729	0.466	0.400	**6.137**	**2**	**0.046**	67		
	Go RT	Male	99	101	8	-0.231	0.144	-0.513	0.050	-1.609	0.108	0.000	3.563	7	0.829	0	0.977	0.329
		Female	140	101	8	-0.035	0.140	-0.309	0.239	-0.254	0.799	0.009	7.415	7	0.387	6		
Stop-signal	SSRT*	Male	176	168	8	-0.206	0.110	-0.421	0.009	-1.879	0.060	0.000	5.119	7	0.645	0	1.345	0.179
		Female	294	132	7	0.053	0.158	-0.257	0.363	0.334	0.731	0.063	9.691	6	0.138	38		
	Go RT	Male	161	152	7	0.029	0.115	-0.196	0.253	0.250	0.803	0.000	4.758	6	0.575	0	-0.289	0.773
		Female	279	117	6	-0.018	0.114	-0.241	0.204	-0.159	0.874	0.000	4.392	5	0.494	0		

Note: Bold type indicates a significant result; positive effect sizes (g) indicate poorer performance by the heavy drinker group compared to controls; k = number of studies; SE = standard error. * The measures of most interest are the NoGo errors (failures of inhibition), and the SSRT (time required to stop a response). In the Go/NoGo task, differences between groups for NoGo errors must be interpreted together with results for Go errors and Go RT (indexing inattention), while in the stop-signal task, SSRT can be interpreted independently of results for Go RT.

Table 5. Characteristics of study samples by behavioural paradigm and study, for papers reporting on alcohol dependent participants vs. controls. Criteria refers to the diagnostic criteria used for entry into control and alcohol dependent groups. Readers are referred to the original papers for further details.

Task/Study	Control group criteria	Alcohol dependent group criteria	Sample size, males		Sample size, females		Group x sex interaction examined?	Notes
			Control	Alcohol dependent	Control	Alcohol dependent		
Go/NoGo task								
Bottesi et al. (2015)	No psychiatric disorder, AUDIT score <5	AUDIT score 5+, detoxified 2+ weeks	37	20	10	20	No	-
Czapla et al. (2016)	No alcohol-related problems, no dependence (except nicotine), psychiatric or neurological disease	DSM-IV AD, no other dependence (except nicotine), psychiatric or neurological diseases, mean abstinence 18.2 days (range 6-76)	54	80	17	20	No	Pooled over alcohol and neutral stimuli
Stein et al. (2018)	Healthy controls without risky drinking habits, AUDIT < 8, no psychopathological symptoms	ICD-10 AD, abstinent 8+ days (mean 28 days), 2 had history of cannabis, cocaine and ecstasy use, 8 were smokers, 6 had no history of other substance use. “At least half” had family history of alcohol problems. Not excluded for other comorbid psychiatric conditions or prescribed medication use	9	12	6	3	No	-
Taylor et al. (2016)	No history of substance dependence (except nicotine)	DSM-IV AD, abstinent 2+ weeks	38	20	14	6	No	-
Stop-signal task								
Choi et al. (2014)	No history of psychiatric disorder	DSM-IV AD, mean AUDIT = 25.5	15	15	-	-	N/A	-
Goudriaan et al. (2006)	No major psychiatric disorders, including AD	DSM-IV AD, abstinent 3-12 months	30	35	18	11	No	-
Hu et al. (2015)	Physically healthy, mean 5.9 occasions/month, 14.1 units/month	DSM-IV AD, no other dependence, residential rehab inpatient. Mean 23.2 occasions/month, 383.6 units/month prior to admission	43	18	27	6	No	-
Lawrence et al. (2009)	No psychiatric illness	DSM-IV AD, 4 subjects had consumed alcohol within 48 hours (but sober at testing); remainder abstinent 1+ week	25	19	-	-	N/A	-
Marin et al. (2015)	No personal or family history of psychiatric disorder	DSM-IV AD, abstinent 1+ month, no other psychiatric disorders	37	40	-	-	N/A	-
Rubio et al. (2007)	No history of psychiatric disorder, no parental suspected drinking problems	DSM-IV AD, detoxified 4-6 weeks	96	247	-	-	N/A	Pooled across personality disorder groups
Schmaal et al. (2013)	No DSM-IV diagnosis except nicotine dependence	DSM-IV AD, no other DSM-IV diagnosis except nicotine dependence	16	16	-	-	N/A	Placebo means used
Sion et al. (2017)	No history of psychiatric problems	Females: 14+ units/week Males: 21+ units/week Attending a hospital alcohol detoxification and recovery program, abstinent 14 days No other psychiatric diagnosis	16	54	7	20	No	Neutral means used
Sjoerds et al. (2013)	No lifetime Axis I diagnoses	DSM-IV AD, no other Axis I diagnosis except depression or anxiety disorder, abstinent 2 weeks	10	17	6	14	No	-
Taylor et al. (2016)	No history of substance dependence (except nicotine)	DSM-IV AD, abstinent 2+ weeks	41	21	13	6	No	-
van der Plas et al. (2009)	No history of psychiatric disorders including substance abuse or dependence	DSM-IV substance dependence (primarily alcohol), not excluded if other drugs used in past 30 days	18	9	16	12	Yes, no significant effects or interactions involving sex	-

Table 6. Summary statistics and weights for each study included in the meta-analysis of Go/NoGo task performance in alcohol dependent groups vs. controls.

	NoGo errors					Go errors					Go RT				
	g	SE	CI lower	CI upper	Weight (%)	g	SE	CI lower	CI upper	Weight (%)	g	SE	CI lower	CI upper	Weight (%)
Males															
Bottesi et al. (2015)	**0.640**	**0.284**	0.084	1.197	19.9	**0.780**	**0.287**	0.217	1.342	39.2	**0.697**	**0.285**	0.138	1.256	25.2
Czapla et al. (2016)	0.332	0.177	-0.016	0.679	51.0						0.155	0.176	-0.191	0.500	33.6
Stein et al. (2018)	0.152	0.442	-0.714	1.017	8.2	0.444	0.446	-0.431	1.318	19.4	0.659	0.453	-0.228	1.546	15.5
Taylor et al. (2016)	0.113	0.276	-0.428	0.655	21.0	0.137	0.277	-0.405	0.679	41.4	-0.328	0.278	-0.873	0.217	25.7
Total	**0.333**	**0.127**	0.084	0.581	100.0	**0.449**	**0.212**	0.034	0.863	100.0	0.245	0.224	-0.193	0.684	100.0
Females															
Bottesi et al. (2015)	-0.062	0.387	-0.821	0.697	28.0	**0.787**	**0.400**	0.002	1.572	53.8	**1.142**	**0.414**	0.330	1.954	27.9
Czapla et al. (2016)	**1.634**	**0.381**	0.888	2.380	28.2						-0.020	0.330	-0.667	0.626	33.1
Stein et al. (2018)	-0.195	0.709	-1.584	1.194	19.2	1.224	0.764	-0.273	2.720	14.8	0.559	0.719	-0.851	1.969	15.1
Taylor et al. (2016)	0.737	0.502	-0.246	1.720	24.7	**1.202**	**0.524**	0.176	2.229	31.4	-0.234	0.489	-1.193	0.725	23.9
Total	0.588	0.463	-0.319	1.494	100.0	**0.982**	**0.294**	0.406	1.557	100.0	0.340	0.336	-0.319	1.000	100.0

Note: Bold type indicates a significant result; positive effect sizes indicate poorer performance by the alcohol dependent group compared to controls

Table 7. Summary statistics and weights for each study included in the meta-analysis of stop-signal task performance in alcohol dependent groups vs. controls.

	SSRT					Go RT				
	g	SE	CI lower	CI upper	Weight (%)	g	SE	CI lower	CI upper	Weight (%)
Males										
Choi et al. (2014)	0.297	0.367	-0.422	1.017	6.8	-0.286	0.367	-1.005	0.433	10.4
Goudriaan et al. (2006)	**0.753**	**0.257**	0.248	1.257	10.2	0.232	0.250	-0.257	0.721	12.7
Hu et al. (2015)	-0.511	0.285	-1.069	0.046	9.2	**0.805**	**0.290**	0.236	1.373	11.9
Lawrence et al. (2009)	0.570	0.310	-0.038	1.179	8.4	**1.242**	**0.332**	0.592	1.893	11.1
Marin et al. (2015)	**0.588**	**0.233**	0.131	1.044	11.2					
Rubio et al. (2007)	**0.334**	**0.121**	0.097	0.572	16.3					
Schmaal et al. (2013)	-0.588	0.361	-1.296	0.120	7.0	-0.333	0.356	-1.031	0.365	10.6
Sion et al. (2017)	0.149	0.285	-0.410	0.707	9.2	-0.161	0.285	-0.720	0.397	12.0
Sjoerds et al. (2013)	0.195	0.399	-0.588	0.978	6.1	0.535	0.405	-0.259	1.329	9.7
Taylor et al. (2016)	0.139	0.269	-0.388	0.665	9.8	-0.072	0.268	-0.598	0.454	12.3
van der Plas et al. (2009)	0.335	0.411	-0.470	1.140	5.8	**-0.967**	**0.429**	-1.807	-0.126	9.3
Total	**0.236**	**0.118**	0.004	0.468	100.0	0.132	0.204	-0.267	0.531	100.0
Females										
Choi et al. (2014)										
Goudriaan et al. (2006)	**0.821**	**0.398**	0.041	1.600	20.5	0.331	0.385	-0.424	1.086	21.3
Hu et al. (2015)	0.131	0.452	-0.755	1.016	15.9	0.399	0.454	-0.491	1.288	15.3
Lawrence et al. (2009)										
Marin et al. (2015)										
Rubio et al. (2007)										
Schmaal et al. (2013)										
Sion et al. (2017)	0.420	0.443	-0.448	1.288	16.5	-0.758	0.451	-1.642	0.126	15.5
Sjoerds et al. (2013)	0.239	0.489	-0.720	1.198	13.5	-0.011	0.488	-0.967	0.946	13.3
Taylor et al. (2016)	0.208	0.495	-0.761	1.178	13.2	-0.302	0.496	-1.274	0.670	12.9
van der Plas et al. (2009)	**0.814**	**0.397**	0.036	1.593	20.5	0.065	0.382	-0.683	0.814	21.7
Total	**0.484**	**0.180**	0.132	0.837	100.0	-0.012	0.178	-0.361	0.337	100.0

Note: Bold type indicates a significant result; positive effect sizes indicate poorer performance by the alcohol dependent group compared to controls

Table 8. Weighted mean effect size (g) and heterogeneity analysis by task, outcome measure, and sex for alcohol dependence vs. controls, as well as statistical tests for sex differences in the weighted mean effect size estimates.

					Effect size analysis							Heterogeneity analysis				Sex differences analysis	
Task	Variable	Sex	Total n, control	Total n, alcohol dependent	k	g	SE	95% CI lower	95% CI upper	z	p	τ^2	Q	df	p	I^2	z	p
Go/NoGo	NoGo errors	Male	138	132	4	**0.333**	0.127	0.084	0.581	**2.628**	**0.009**	0.000	1.972	3	0.578	0	0.532	0.595
		Female	47	49	4	0.588	0.463	-0.319	1.494	1.270	0.204	0.615	**11.489**	**3**	**0.009**	74		
	Go errors	Male	84	52	3	**0.449**	0.212	0.034	0.863	**2.119**	**0.034**	0.032	2.601	2	0.272	23	1.474	0.140
		Female	30	29	3	**0.982**	0.294	0.406	1.557	**3.344**	**0.001**	0.000	0.515	2	0.773	0		
	Go RT	Male	138	132	4	0.245	0.224	-0.193	0.684	1.096	0.273	0.118	7.741	3	0.052	61	0.235	0.814
		Female	47	49	4	0.340	0.336	-0.319	1.000	1.012	0.312	0.234	6.389	3	0.094	53		
Stop-signal	SSRT	Male	347	491	11	**0.236**	0.118	0.004	0.468	**1.991**	**0.046**	0.071	**20.265**	**10**	**0.027**	51	1.155	0.248
		Female	87	69	6	**0.484**	0.180	0.132	0.837	**2.694**	**0.007**	0.000	2.604	5	0.761	0		
	Go RT	Male	214	204	9	0.132	0.204	-0.267	0.531	0.648	0.517	0.264	**28.824**	**8**	**0.000**	72	-0.533	0.594
		Female	87	69	6	-0.012	0.178	-0.361	0.337	-0.067	0.947	0.000	4.728	5	0.450	0		

Note: Bold type indicates a significant result; positive effect sizes (g) indicate poorer performance by the alcohol dependent group compared to controls; k = number of studies; SE = standard error.

4. DISCUSSION

In this report, we focus on the evidence for sex differences in the relationship between heavy alcohol consumption and poorer behavioural inhibition. In the discussion below, we review (a) our results in comparison with previous work, and possible reasons for differences in results; (b) the evidence concerning the existence of sex differences in the relationship between heavy alcohol use and disinhibition for heavy drinkers and alcohol dependent groups; (c) possible sources of heterogeneity; and (d) suggestions for future research and the implications of the current results for assessment and treatment.

4.1 Comparison with previous work

In comparison to our previous analysis (Smith et al., 2014), which estimated the effect size for inhibitory dysfunction in heavy drinkers and alcohol dependent groups regardless of sex, we observed some similarities and some differences in our current results. For alcohol dependence, we previously observed significant reductions in inhibitory capacity in the frequent-Go/rare-NoGo task (g = 0.531) and the stop signal task (g = 0.395); these results were matched in the current study with small-medium sized but significant effects for both men and women for SSRT, and significant small deficits for alcohol dependent men for NoGo errors (see Table 8). Further, this updated meta-analysis has provided new information: previously, we were unable to calculate weighted mean effect sizes for Go errors and Go RT because too few studies had reported these; here, confirmation of an increase in Go errors for alcohol dependent men and women relative to controls indicates a problem with discrimination of Go and NoGo stimuli, rather than a problem with inhibition specifically.

In contrast, the small but significant effect for heavy drinkers in the stop-signal task (g = 0.248; Smith et al., 2014) was not replicated here (males: g = -0.206, females: g = 0.053). These differences between the 2014 results and the current results can be explained by the different studies included in each analysis. Several new papers have been published since the previous analyses were performed, and have been included here (Bø and Landrø, 2017, Hu et al., 2016, Kareken et al., 2013, Smith et al., 2016, van Duijvenbode et al., 2013); similarly, some papers that were included in the 2014 analysis were not included here due to the stricter requirements on NoGo probability and maximum SOA, or we were unable to access sufficient information for calculating effect sizes individually for each sex. This includes some papers with very large sample sizes and/or large effect sizes between heavy drinkers and controls (e.g., Moreno et al., 2012, Rubio et al., 2008); these exclusions alter the size and precision of the weighted mean effect estimate.

4.2 Evidence of sex differences

The expected sex differences failed to materialise for heavy drinkers; not only did the statistical comparison of effect sizes fail to reach significance, but there was not even any clear pattern of more positive effects for women compared to men. In fact, for the measure which showed the largest difference in effect sizes between men and women (SSRT), the results indicated slightly *better* performance among heavy drinking males, not *poorer* performance among heavy drinking females. There are several possible reasons for this failure to observe the expected reduction in inhibitory control in heavy drinking women but not men. As we mentioned, the effect size estimations depend on the data included in the analysis; we could not access data for some studies which report a significant sex difference, as well as studies which report no significant group × sex interaction (Sanchez-Roige et al., 2014). Thus, the estimates of effect size could be expected to change if more/different data were included. Low statistical power is also likely a problem; most studies were powered only to detect group effects but not group × sex interactions. However, the weighted mean effect sizes are generally small, suggesting a lack of effect rather than a lack of power. That is, the most probable explanation for the lack of significant differences in female but not male heavy drinkers is that there really is no sex difference in the relationship between heavy drinking and behavioural disinhibition, and that the papers reporting a significant group × sex interaction

may be simply false positives. The implication of the lack of observable sex differences is that the same treatment and prevention strategies can be aimed at men and women.

We cannot make strong claims about the presence or absence of sex differences in inhibitory capacity associated with alcohol dependence. On one hand, alcohol-related problems with inhibition and stimulus discrimination were generally the same for both men and women, as were intact response execution times, measured by Go RT in each task; sex differences were not statistically significant for any measure. On the other hand, the effect sizes for women were generally larger than those for men for both NoGo and Go errors, suggestive of greater problems in alcohol dependent women. Because the total sample size for women was very small, caution is necessary for interpreting results. It is possible that we do not have adequate power to detect a true sex difference; it is also possible that the effect sizes for women are falsely inflated, since low power decreases the probability that a given statistically significant result represents a true effect (Button et al., 2013). In the absence of clear evidence for or against the notion of sex differences in the relationship between disinhibition and alcohol dependence, we suggest that more investigation of inhibitory problems in alcohol dependent women is warranted. It is also apparent that studies which seek to examine inhibitory control in alcohol dependence, but which sample only men (or predominantly men), may be underestimating the decrement, and that a failure to adequately sample women and test for sex effects has so far missed a possible opportunity to identify a sex-sensitive effect (Wetherington, 2007). If future research confirms sex differences in inhibitory control among alcohol dependent men and women, potential impacts include the possibility of different treatment strategies for men and women, and future research could investigate when such sex differences arise, since they are not apparent among younger heavy drinkers.

4.3 Heterogeneity

There are several possible sources of heterogeneity of effect sizes between studies, some of which relate to task design, and others to group characterisation. While our Go/NoGo task criteria (Go probability at least 51% and mean SOA ≤ 2000ms) was designed to include only studies with a homogenous design known to tax inhibitory control (Wessel, 2018), considerable heterogeneity remains in other task design parameters. Some studies used neutral stimuli such as letters or geometric shapes in the Go/NoGo task (Czapla et al., 2015), while others used alcohol-related stimuli (Petit et al., 2012); studies have suggested that inhibition is more difficult when alcohol-related stimuli are presented (Kreusch et al., 2013), but others report equivalent inhibitory control for alcohol and neutral stimuli (Nederkoorn et al., 2009). Some studies provide monetary reward for good performance (Rossiter et al., 2012), while most do not. Some tasks involved working memory in addition to inhibitory processes (Smith et al., 2017), while most did not. Such methodological differences between studies may contribute to heterogeneity in the effect sizes observed, since some studies may tax inhibitory capacity more than others, but there were too few studies to conduct a formal analysis of these potential moderators. Investigation of these parameters may prove a fruitful avenue for future research.

Heterogeneity was also observed between studies in group characterisation. Studies differ in the extent to which polydrug use (especially nicotine), comorbid psychopathologies, and family history of alcohol or other psychiatric problems were assessed and excluded or included, and in the length of abstinence for alcohol dependent groups. We note that many of these factors may themselves be associated with inhibitory dysfunction; for example, inhibitory problems are also apparent for attention-deficit/hyperactivity disorder, anxiety disorder, obsessive-compulsive disorder, reading disorder and schizophrenia (Lipszyc and Schachar, 2010), depression (Snyder, 2013) and traumatic brain injury (Dimoska-Di Marco et al., 2011). Substance abuse is commonly comorbid with these and other psychiatric disorders (Petrakis et al., 2002, McKetin et al., 2005, Kaye and Darke, 2004), such that one cannot be certain, given the current evidence, that observed deficits are associated with alcohol dependence and not other comorbid psychopathologies. Similarly, poorer inhibitory control is reported for participants with a family history of substance problems, compared to those with no such family history, even when participants are matched for their own personal use (Acheson et al., 2011) although not all studies report this effect (Saunders et al., 2008). Because polydrug use, comorbid psychopathology and/or family history of substance abuse

are all widespread among people who abuse drugs, we urge researchers in future to assess and report such information, perhaps presenting data separately for “pure” vs. “typical” users, where numbers allow.

For alcohol dependence, studies were reasonably consistent in requiring a DSM-IV or ICD-10 alcohol dependence diagnosis, while for heavy drinkers, there was considerable variation in the measures used to define groups. For example, studies defined groups according to the frequency of binge drinking, total drinks consumed per unit of time, or score on the Alcohol Use Disorders Identification Test (AUDIT: Babor et al., 2001); studies differed on whether abstainers were eligible for inclusion in the control group, and even in the cut-offs used for a single measure (e.g., AUDIT score above 8: Weafer et al., 2015; 12 or more: Petit et al., 2012; or 16 or more: van Duijvenbode et al., 2013; see Table 1). So that the reader may assess the heterogeneity between studies for themselves, we have included not only the minimum criteria for entry to the study, but also the mean for the AUDIT and grams of ethanol consumed/week, where available. It is apparent that there are marked differences between studies in mean use, for both control and heavy drinker groups. For example, the mean AUDIT score for controls (7.1) in Nederkoorn et al. (2009) suggests that at least some of their sample may have scored above the threshold for hazardous and harmful drinking (8 or above: Babor et al., 2001), while the mean AUDIT for heavy drinkers in Bednarski et al. (2012: 7.0) and Hu et al. (2016: 4.9) were below this threshold. Similar comments can be made concerning grams ethanol consumed per week for these studies. However, we have confirmed that the pattern of results does not change substantially when these papers are excluded from analysis, suggesting that this heterogeneity in alcohol consumption between studies does not alter the results observed (or lack thereof). However, it will be important in future research to collect and report as much information as possible to characterise both control and heavy drinker groups, and indeed, move toward a consensus on control and heavy drinker group construction.

Lastly, there was also variation in whether identical or different criteria were applied for males and females; the majority of studies used identical criteria, but several had higher thresholds for categorisation to the heavy drinker group for men than women (Bednarski et al., 2012, Nederkoorn et al., 2009, Watson et al., 2016). Given that the major health organisations differ on separate vs. identical thresholds for defining binge drinking in men and women (NIAAA, 2004, WHO, 2014), we included studies with both approaches, but allow that this might also contribute to heterogeneity.

4.4 Suggestions for future research

The most obvious issue that needs attention in future research is sampling more women in studies of alcohol dependence vs. controls, and including sex as a factor in these analyses. It is possible that at least some of the studies which recruited both men and women included sex as a factor in preliminary analyses, but on finding no significant sex differences, did not mention these analyses in their publications. The greater problem is the absence of any female participants in close to half of the studies of alcohol dependence, and under-sampling of women in the remaining studies, a problem which has been reported elsewhere in substance abuse research (Lind et al., 2017). The under-sampling of women cannot be justified by the apparently greater lifetime prevalence of substance abuse in males (Compton et al., 2007), since evidence (reviewed in the Introduction) suggests a more severe disorder in women. Rather, the under-sampling of women demonstrated here is reminiscent of the historically noted sex bias apparent across many fields of medical research, including animal studies (Beery and Zucker, 2011, Berkley, 1992, Wizemann and Pardue, 2001), and already recognised as in need of correction (Clayton and Collins, 2014, Wetherington, 2007). The United States National Institutes of Health (NIH) and the Food and Drug Administration (FDA) have taken steps to increase the representation of women (and female animals and cells) in clinical research (USFDA, 1993, Clayton and Collins, 2014, NIH, 2001), and provide excellent guidelines not only for researchers, but also funding agencies, journal editors and reviewers, to ensure health research is inclusive and representative (Beery and Zucker, 2011, Hayes and Redberg, 2008, Wizemann and Pardue, 2001). However, we point out that all the studies included in this meta-analysis were published well after the 2001 NIH guidelines, suggesting that despite the NIH advice, under-sampling of women appears to be a persistent problem.

Significant barriers to research participation continue to be faced disproportionately by women, including conflicts between family/childcare obligations and the time required for research participation (Liu and Mager, 2016). In future, researchers will need to consider these guidelines in designing studies with sufficient power for appropriate assessment of sex differences, and minimise or remove barriers to participation for women.

One historical justification for the focus on male participants has been the view that fluctuating hormones associated with the female reproductive cycle introduce additional variability into the data. It is necessary to take these variations into account when studying sex differences in a trait (Becker et al., 2005), particularly since inhibitory control is one such trait that varies according to the menstrual cycle. Healthy control females have significantly longer SSRT in the follicular phase, but do not differ from men during the luteal phase or menstruation proper; increased oestradiol levels are associated with longer SSRT; and response execution as measured by Go RT does not differ between menstrual phases (Colzato et al., 2010). Those authors point out that when menstrual phase is unrecorded, the observed sex differences (or lack thereof) may be due to the exact ratio of women in each phase of the cycle. It is unlikely to be feasible (and is likely outside the scope of the research question) for most studies of inhibitory control in substance abusers to recruit sufficient women in each phase for meaningful sub-group comparisons (Becker et al., 2005), but publication of contraceptive use, menopausal status, and hormonal status (ideally, assessed with biological measures such as saliva or blood hormone levels, rather than self-report: Becker et al., 2005, Wizemann and Pardue, 2001) would allow more easy assessment of menstrual phase as a contributor to heterogeneity within and between studies of inhibitory control in substance-using populations.

Like many of the studies we refer to in this review, we have used the terms ‘sex’ and ‘gender’ interchangeably, when in fact it is unclear whether the observed ‘sex’ differences are due to sex (biological) or gender (psycho-socio-cultural) factors. The current study cannot, and was not designed to, address this important question, nor to address the possible mechanisms behind the observed sex differences. Given the body of evidence suggests that disinhibition pre-dates the onset of substance abuse (Mahmood et al., 2013, Norman et al., 2011, Tarter et al., 2003), it seems likely that pre-existing individual differences in impulsivity/inhibitory capacity may be a greater risk factor for the subsequent development of alcohol abuse in women than men (Moeller et al., 2016). As Nederkoorn et al. (2009) hypothesise, among young males, heavy drinking is prevalent and likely to develop due to social norms, regardless of individual impulsivity; in contrast, among young females, heavy drinking is less prevalent, and only the most disinhibited young women commence heavy drinking. However, this is speculative, and it remains for future longitudinal studies to determine the cause and effect nature of the relationship between inhibitory dysfunction and substance use, and its possible interaction with sex and gender, and to assess the effectiveness of sex-based early intervention programs targeted at remediating inhibitory differences to improve future outcomes (Moeller et al., 2016).

In this meta-analysis, we focused on alcohol because there were several previous studies reporting sex differences in inhibitory control among heavy drinkers (Nederkoorn et al., 2009, Townshend and Duka, 2005, Smith et al., 2016, Kreusch et al., 2013, Weafer et al., 2015), because our previous meta-analysis (Smith et al., 2014) identified studies on alcohol dependence as being particularly skewed to oversampling males, and because there were a sufficient number of studies of heavy drinkers and alcohol dependent groups to allow a meta-analytic approach. However, it remains to be seen whether sex differences are observed in users of other substances; there is some suggestion in previous research that response inhibition is better among male smokers than male non-smokers (Fields et al., 2009, Reynolds et al., 2007), although poorer inhibition among female smokers was observed only in Fields et al. (2009) and not by Reynolds et al. (2007). Therefore, examination of sex differences in inhibitory dysfunction should be an aim of future research on abuse of all substances, not merely limited to alcohol.

4.5 Conclusion

In conclusion, this meta-analytic review of 31 studies has examined the evidence for sex differences in inhibitory dysfunction in heavy drinkers and alcohol dependent groups. Despite some previous reports of greater inhibitory dysfunction in female heavy drinkers, there was no meta-analytic evidence of such a sex difference. For alcohol dependence, despite minimal investigation of sex differences in previous research, there was some suggestion of greater dysfunction among women, although the sex differences did not reach statistical significance; low statistical power suggests that caution is required in interpreting this result. The main limitation of the current study is that the number of studies examining this issue is small, and most of these studies contained very few participants, especially so for women. The results highlight the need to include more women in research, with sufficient sample sizes to allow analysis of sex differences, and particularly in relation to the contribution of inhibitory dysfunction to the development and maintenance of substance dependence in women.

5. REFERENCES

Acheson, A., Richard, D.M., Mathias, C.W., & Dougherty, D.M. (2011) Adults with a family history of alcohol related problems are more impulsive on measures of response initiation and response inhibition. *Drug and Alcohol Dependence* 117, 198-203.

Allom, V., Mullan, B., & Hagger, M. (2016) Does inhibitory control training improve health behaviour? A meta-analysis. *Health Psychology Review* 10, 168-186.

Anthony, J.C., Warner, L.A., & Kessler, R.C. (1994) Comparative epidemiology of dependence on tobacco, alcohol, controlled substances, and inhalants: Basic findings from the National Comorbidity Survey. *Experimental and Clinical Psychopharmacology* 2, 244-268.

Babor, T.F., Higgins-Biddle, J.C., Saunders, J.B., & Monteiro, M.G. (2001) AUDIT: the Alcohol Use Disorders Identification Test: guidelines for use in primary health care, Geneva, World Health Organization.

Band, G.P.H., Van der Molen, M.W., & Logan, G.D. (2003) Horse-race model simulations of the stop-signal procedure. *Acta Psychologica* 112, 105-142.

Barkley, R.A. (1997) Behavioural inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. *Psychological Bulletin* 121, 65-94.

Becker, J.B., Arnold, A.P., Berkley, K.J., Blaustein, J.D., Eckel, L.A., Hampson, E., Herman, J.P., Marts, S., Sadee, W., Steiner, M., Taylor, J., & Young, E. (2005) Strategies and methods for research on sex differences in brain and behavior. *Endocrinology* 146, 1650-1673.

Bednarski, S.R., Erdman, E., Luo, X., Zhang, S., Hu, S., & Li, C.-S.R. (2012) Neural processes of an indirect analog of risk taking in young nondependent adult alcohol drinkers—An fMRI study of the stop signal task. *Alcoholism: Clinical and Experimental Research* 36, 768-779.

Beery, A.K., & Zucker, I. (2011) Sex bias in neuroscience and biomedical research. *Neuroscience and Biobehavioral Reviews* 35, 565-572.

Berkley, K.J. (1992) Vive la différence! *Trends in Neurosciences* 15, 331-332.

Bø, R., & Landrø, N.I. (2017) Inhibitory control and response monitoring are not systematically related to weekly alcohol consumption in the general population. *Psychopharmacology* 1-8.

Borenstein, M., & Higgins, J.P.T. (2013) Meta-analysis and subgroups. *Prevention Science* 14, 134-143.

Bottesi, G., Ghisi, M., Ouimet, A.J., Tira, M.D., & Sanavio, E. (2015) Compulsivity and impulsivity in pathological gambling: Does a dimensional–transdiagnostic approach add clinical utility to DSM-5 classification? *Journal of Gambling Studies* 31, 825-847.

Bruin, K.J., & Wijers, A.A. (2002) Inhibition, response mode, and stimulus probability: a comparative event-related potential study. *Clinical Neurophysiology* 113, 1172-1182.

Button, K.S., Ioannidis, J.P.A., Mokrysz, C. Nosek, B.A., Flint, J., Robinson, E.S.J., & Munafo, M.R. (2013) Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience* 14, 365-376.

Campanella, S., Absil, J., Carbia Sinde, C., Schroder, E., Peigneux, P., Bourguignon, M., Petieau, M., Metens, T., Nouali, M., Goldman, S., Cheron, G., Verbanck, P., & De Tiège, X. (2017) Neural correlates of correct and failed response inhibition in heavy versus light social

drinkers: an fMRI study during a go/no-go task by healthy participants. *Brain Imaging and Behavior* 11*,* 1796-1811.

Choi, S.-W., Kim, H.S., Kim, G.-Y., Jeon, Y., Park, S.M., Lee, J.-Y., Jung, H.Y., Sohn, B.K., Choi, J.-S., & Kim, D.-J. (2014) Similarities and differences among Internet gaming disorder, gambling disorder and alcohol use disorder: A focus on impulsivity and compulsivity. *Journal of Behavioral Addictions* 3, 246-253.

Clayton, J.A., & Collins, F.S. (2014) NIH to balance sex in cell and animal studies. *Nature* 509, 282-283.

Cohen, J. (1992) A power primer. *Psychological Bulletin* 112, 155-159.

Colzato, L.S., Hertsig, G., van den Wildenberg, W.P.M., & Hommel, B. (2010) Estrogen modulates inhibitory control in healthy human females: evidence from the stop-signal paradigm. *Neuroscience* 167, 709-715.

Compton, W.M., Thomas, Y.F., Stinson, F.S., & Grant, B.F. (2007) Prevalence, correlates, disability, and comorbidity of DSM-IV drug abuse and dependence in the United States: results from the national epidemiologic survey on alcohol and related conditions. *Archives of General Psychiatry* 64, 566-576.

Czapla, M., Simon, J.J., Friederich, H.C., Herpertz, S.C., Zimmermann, P., & Loeber, S. (2015) Is binge drinking in young adults associated with an alcohol-specific impairment of response inhibition. *European Addiction Research* 21, 105-113.

Czapla, M., Simon, J.J., Richter, B., Kluge, M., Friederich, H.-C., Herpertz, S., Mann, K., Herpertz, S.C., & Loeber, S. (2016) The impact of cognitive impairment and impulsivity on relapse of alcohol-dependent patients: implications for psychotherapeutic treatment. *Addiction Biology* 21, 873-884.

Dimoska-Di Marco, A., McDonald, S., Kelly, M., Tate, R., & Johnstone, S. (2011) A meta-analysis of response inhibition and Stroop interference control deficits in adults with traumatic brain injury (TBI). *Journal of Clinical and Experimental Neuropsychology* 33, 471-485.

Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997) Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal* 315, 629-634.

Erol, A., & Karpyak, V.M. (2015) Sex and gender-related differences in alcohol use and its consequences: Contemporary knowledge and future research considerations. *Drug and Alcohol Dependence* 156, 1-13.

Fields, S., Collins, C., Leraas, K., & Reynolds, B. (2009) Dimensions of impulsive behavior in adolescent smokers and nonsmokers. *Experimental and Clinical Psychopharmacology* 17, 302-311.

Franken, I.H.A., Luijten, M., van der Veen, F.M., & van Strien, J.W. (2017) Cognitive control in young heavy drinkers: An ERP study. *Drug and Alcohol Dependence* 175, 77-83.

Goldstein, R.Z., & Volkow, N.D. (2002) Drug addiction and its underlying neurobiological basis: Neuroimaging evidence for the involvement of the frontal cortex. *American Journal of Psychiatry* 159, 1642-1652.

Goudriaan, A.E., Oosterlaan, J., De Beurs, E., & Van Den Brink, W. (2006) Neurocognitive functions in pathological gambling: a comparison with alcohol dependence, Tourette syndrome and normal controls. *Addiction* 101, 534-547.

Goudriaan, A.E., Oosterlaan, J., De Beurs, E., & Van Den Brink, W. (2008) The role of self-reported impulsivity and reward sensitivity versus neurocognitive measures of disinhibition

and decision-making in the prediction of relapse in pathological gamblers. *Psychological Medicine* 38, 41-50.

Hayes, S.N., & Redberg, R.F. (2008) Dispelling the myths: Calling for sex-specific reporting of trial results. *Mayo Clinic Proceedings* 83, 523-525.

Hedges, L.V. (1980) Unbiased estimation of effect size. *Evaluation in Education* 4, 25-27.

Hedges, L.V., & Olkin, I. (1985) Statistical methods for meta-analysis, Orlando, Florida, Academic Press.

Hester, R., Lubman, D.I., & Yücel, M. (2010) The role of executive control in human drug addiction. In Self, D.W. & Staley Gottschalk, J.K. (Eds.) Behavioral Neuroscience of Drug Addiction. Springer Berlin Heidelberg.

Higgins, J.P.T., Thompson, S.G., Deeks, J.J., & Altman, D.G. (2003) Measuring inconsistency in meta-analyses. *British Medical Journal* 327, 557-560.

Hommer, D.W. (2003) Male and female sensitivity to alcohol-induced brain damage. *Alcohol Research & Health* 27, 181-185.

Hser, Y.-I., Evans, E., & Huang, Y.-C. (2005) Treatment outcomes among women and men methamphetamine abusers in California. *Journal of Substance Abuse Treatment* 28, 77-85.

Hu, S., Ide, J.S., Zhang, S., Sinha, R., & Li, C.-S.R. (2015) Conflict anticipation in alcohol dependence — A model-based fMRI study of stop signal task. *NeuroImage: Clinical* 8, 39-50.

Hu, S., Zhang, S., Chao, H.H., Krystal, J.H., & Li, C.-S.R. (2016) Association of drinking problems and duration of alcohol use to inhibitory control in nondependent young adult social drinkers. *Alcoholism: Clinical and Experimental Research* 40, 319-328.

Jones, A., Di Lemma, L.C.G., Robinson, E., Christiansen, P., Nolan, S., Tudur-Smith, C., & Field, M. (2016) Inhibitory control training for appetitive behaviour change: A meta-analytic investigation of mechanisms of action and moderators of effectiveness. *Appetite* 97, 16-28.

Kareken, D.A., Dzemidzic, M., Wetherill, L., Eiler, W., Oberlin, B.G., Harezlak, J., Wang, Y., & O'Connor, S.J. (2013) Family history of alcoholism interacts with alcohol to affect brain regions involved in behavioral inhibition. *Psychopharmacology* 228, 335-345.

Kaye, S., & Darke, S. (2004) Injecting and non-injecting cocaine use in Australia: physical and psychological morbidity. *Drug and Alcohol Review* 23, 391-398.

Kreusch, F., Quertemont, E., Vilenne, A., & Hansenne, M. (2014) Alcohol abuse and ERP components in Go/No-go tasks using alcohol-related stimuli: Impact of alcohol avoidance. *International Journal of Psychophysiology* 94, 92-99.

Kreusch, F., Vilenne, A.l., & Quertemont, E. (2013) Response inhibition toward alcohol-related cues using an alcohol go/no-go task in problem and non-problem drinkers. *Addictive Behaviors* 38, 2520-2528.

Lawrence, A., Luty, J., Bogdan, N., Sahakian, B., & Clark, L. (2009) Impulsivity and response inhibition in alcohol dependence and problem gambling. *Psychopharmacology* 207, 163-172.

Lind, K.E., Gutierrez, E.J., Yamamoto, D.J., Regner, M.F., McKee, S.A., & Tanabe, J. (2017) Sex disparities in substance abuse research: Evaluating 23 years of structural neuroimaging studies. *Drug and Alcohol Dependence* 173, 92-98.

Lipsey, M.W., & Wilson, D.B. (2001) Practical Meta-analysis, California, Sage Publications.

Lipszyc, J., & Schachar, R. (2010) Inhibitory control and psychopathology: A meta-analysis of studies using the stop signal task. *Journal of the International Neuropsychological Society* 16, 1064-1076.

Liu, K.A., & Mager, N.A.D. (2016) Women's involvement in clinical trials: historical perspective and future implications. *Pharmacy Practice* 14, 708.

Logan, G.D. (1994) On the ability to inhibit thought and action: A users' guide to the stop signal paradigm. In Dagenbach, D. & Carr, T.H. (Eds.) Inhibitory processes in attention, memory and language. San Diego, California, Academic Press.

Logan, G.D., & Cowan, W.B. (1984) On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review* 91, 295-327.

Luijten, M., Kleinjan, M., & Franken, I.H. (2016) Event-related potentials reflecting smoking cue reactivity and cognitive control as predictors of smoking relapse and resumption. *Psychopharmacology* 233, 2857-2868.

Mahmood, O.M., Goldenberg, D., Thayer, R., Migliorini, R., Simmons, A.N., & Tapert, S.F. (2013) Adolescents' fMRI activation to a response inhibition task predicts future substance use. *Addictive Behaviors* 38, 1435-1441.

Mann, K., Ackermann, K., Croissant, B., Mundle, G., Nakovics, H., & Diehl, A. (2005) Neuroimaging of gender differences in alcohol dependence: Are women more vulnerable? *Alcoholism: Clinical and Experimental Research* 29, 896-901.

Marín-Martínez, F., & Sánchez-Meca, J. (1998) Testing for dichotomous moderators in meta-analysis. *The Journal of Experimental Education* 67, 69-81.

Marin, M., Jurado, R., Ponce, G., Koeneke, A., Martinez-Gras, I., Jiménez-Arriero, M.Á., & Rubio, G. (2015) Startle-response based tasks and laboratory measures of impulsivity in abstinent alcoholic patients. *Alcohol and Alcoholism* 50, 286-295.

McKetin, R., McLaren, J., & Kelly, E. (2005) The Sydney methamphetamine market: patterns of supply, use, personal harms and social consequences. National Drug Law Enforcement Research Fund Monograph No 13. Adelaide, Australasian Centre for Policing Studies.

Mereu, M., Bonci, A., Newman, A., & Tanda, G. (2013) The neurobiology of modafinil as an enhancer of cognitive performance and a potential treatment for substance use disorders. *Psychopharmacology* 229, 415-434.

Moeller, S.J., Bederson, L., Alia-Klein, N., & Goldstein, R.Z. (2016) Neuroscience of inhibition for addiction medicine: from prediction of initiation to prediction of relapse. In Hamed, E. & Martin, P. (Eds.) Progress in Brain Research. Elsevier.

Moreno, M., Estevez, A.F., Zaldivar, F., Montes, J.M.G., Gutiérrez-Ferre, V.E., Esteban, L., Sánchez-Santed, F., & Flores, P. (2012) Impulsivity differences in recreational cannabis users and binge drinkers in a university population. *Drug and Alcohol Dependence* 124, 355-362.

National Institute on Alcohol Abuse and Alcoholism (NIAAA) (2004) NIAAA approves definition of binge drinking. Newsletter 3 (Winter).

National Institutes of Health (NIH) (2001) NIH policy and guidelines on the inclusion of women and minorities as subjects in clinical research.

Nederkoorn, C., Baltus, M., Guerrieri, R., & Wiers, R.W. (2009) Heavy drinking is associated with deficient response inhibition in women but not in men. *Pharmacology Biochemistry and Behavior* 93, 331-336.

Nieuwenhuis, S., Yeung, N., van den Wildenberg, W., & Ridderinkhof, K.R. (2003) Electrophysiological correlates of anterior cingulate function in a go/no-go task: Effects of response conflict and trial type frequency. *Cognitive, Affective, & Behavioral Neuroscience* 3, 17-26.

Norman, A.L., Pulido, C., Squeglia, L.M., Spadoni, A.D., Paulus, M.P., & Tapert, S.F. (2011) Neural activation during inhibition predicts initiation of substance use in adolescence. *Drug and Alcohol Dependence* 119, 216-223.

Papachristou, H., Nederkoorn, C., Havermans, R., van der Horst, M., & Jansen, A. (2012) Can't stop the craving: The effect of impulsivity on cue-elicited craving for alcohol in heavy and light social drinkers. *Psychopharmacology* 219, 511-518.

Petit, G., Cimochowska, A., Kornreich, C., Hanak, C., Verbanck, P., & Campanella, S. (2014) Neurophysiological correlates of response inhibition predict relapse in detoxified alcoholic patients: Some preliminary evidence from event-related potentials. *Neuropsychiatric Disease and Treatment* 10, 1025-1037.

Petit, G., Kornreich, C., Noël, X., Verbanck, P., & Campanella, S. (2012) Alcohol-related context modulates performance of social drinkers in a visual Go/No-Go task: a preliminary assessment of event-related potentials. *PLoS ONE* 7, e37466.

Petrakis, I.L., Gonzalez, G., Rosenheck, R., & Krystal, J.H. (2002) Comorbidity of alcoholism and psychiatric disorders. *Alcohol Research & Health* 26, 81-89.

Piazza, N.J., Vrbka, J.L., & Yeager, R.D. (1989) Telescoping of alcoholism in women alcoholics. *International Journal of the Addictions* 24, 19-28.

Prisciandaro, J.J., Myrick, H., Henderson, S., McRae-Clark, A.L., & Brady, K.T. (2013) Prospective associations between brain activation to cocaine and no-go cues and cocaine relapse. *Drug and Alcohol Dependence* 131, 44-49.

Randall, W.M., & Smith, J.L. (2011) Conflict and inhibition in the cued-Go/NoGo task. *Clinical Neurophysiology* 122, 2400-2407.

Reynolds, B., Patak, M., Shroff, P., Penfold, R.B., Melanko, S., & Duhig, A.M. (2007) Laboratory and self-report assessments of impulsive behavior in adolescent daily smokers and nonsmokers. *Experimental and Clinical Psychopharmacology* 15, 264-271.

Rossiter, S., Thompson, J., & Hester, R. (2012) Improving control over the impulse for reward: Sensitivity of harmful alcohol drinkers to delayed reward but not immediate punishment. *Drug and Alcohol Dependence* 125, 89-94.

Rubio, G., Jiménez, M., Rodríguez-Jiménez, R., Martínez, I., Ávila, C., Ferre, F., Jiménez-Arriero, M.A., Ponce, G., & Palomo, T. (2008) The role of behavioral impulsivity in the development of alcohol dependence: A 4-year follow-up study. *Alcoholism: Clinical and Experimental Research* 32, 1681-1687.

Rubio, G., Jiménez, M., Rodríguez-Jiménez, R., Martínez, I., Iribarren, M.M., Jiménez-Arriero, M.A., Ponce, G., & Ávila, C. (2007) Varieties of impulsivity in males with alcohol dependence: The role of Cluster-B Personality Disorder. Alcoholism: Clinical and Experimental Research, 31, 1826-1832.

Rupp, C.I., Beck, J.K., Heinz, A., Kemmler, G., Manz, S., Tempel, K., & Fleischhacker, W.W. (2016) Impulsivity and alcohol dependence treatment completion: Is there a neurocognitive risk factor at treatment entry? *Alcoholism: Clinical and Experimental Research* 40, 152-160.

Sánchez-Meca, J., & Marín-Martínez, F. (1998) Testing continuous moderators in meta-analysis: A comparison of procedures. *British Journal of Mathematical and Statistical Psychology* 51, 311-326.

Sanchez-Roige, S., Baro, V., Trick, L., Peña-Oliver, Y., Stephens, D.N., & Duka, T. (2014) Exaggerated waiting impulsivity associated with human binge drinking, and high alcohol consumption in mice. *Neuropsychopharmacology* 39, 2919-2927.

Saunders, B., Farag, N., Vincent, A.S., Frank L. Collins, J., Sorocco, K.H., & Lovallo, W.R. (2008) Impulsive errors on a Go-NoGo reaction time task: Disinhibitory traits in relation to a family history of alcoholism. *Alcoholism: Clinical and Experimental Research* 32, 888-894.

Schmaal, L., Joos, L., Koeleman, M., Veltman, D.J., van den Brink, W., & Goudriaan, A.E. (2013) Effects of modafinil on neural correlates of response inhibition in alcohol-dependent patients. *Biological Psychiatry* 73, 211-218.

Sharrett-Field, L., Butler, T.R., Reynolds, A.R., Berry, J.N., & Prendergast, M.A. (2013) Sex differences in neuroadaptation to alcohol and withdrawal neurotoxicity. *Pflugers Archiv: European Journal of Physiology* 465, 643-654.

Sion, A., Jurado-Barba, R., Alonso, M.J., & Rubio-Valladolid, G. (2017) Inhibitory capacity assessment in alcohol dependent patients: translation from a modified stop signal task. *Actas Espanolas de Psiquiatria* 45, 21-31.

Sjoerds, Z., van den Brink, W., Beekman, A.T.F., Penninx, B.W.J.H., & Veltman, D.J. (2014) Response inhibition in alcohol-dependent patients and patients with depression/anxiety: a functional magnetic resonance imaging study. *Psychological Medicine* 44, 1713-1725.

Slade, T., Chapman, C., Swift, W., Keyes, K., Tonks, Z., & Teesson, M. (2016) Birth cohort trends in the global epidemiology of alcohol use and alcohol-related harms in men and women: systematic review and metaregression. *BMJ Open* 6, e011827.

Smith, J.L., Iredale, J.M., & Mattick, R.P. (2016) Sex differences in the relationship between heavy alcohol use, inhibition and performance monitoring: disconnect between behavioural and brain functional measures. *Psychiatry Research: Neuroimaging* 254, 103-111.

Smith, J.L., Mattick, R.P., Jamadar, S., & Iredale, J.M. (2014) Deficits in behavioural inhibition in substance abuse and addiction: A meta-analysis. *Drug and Alcohol Dependence* 145, 1-33.

Smith, J.L., Mattick, R.P., & Sufani, C. (2017) Error detection and behavioural inhibition in young heavy drinkers. *Drug and Alcohol Dependence* 171, 20-30.

Snyder, H.R. (2013) Major depressive disorder is associated with broad impairments on neuropsychological measures of executive function: A meta-analysis and review. *Psychological Bulletin* 139, 81-132.

Spechler, P.A., Chaarani, B., Hudson, K.E., Potter, A., Foxe, J.J., & Garavan, H. (2016) Response inhibition and addiction medicine: from use to abstinence. In Hamed, E. & Martin, P. (Eds.) Progress in Brain Research. Elsevier.

Steele, V.R., Fink, B.C., Maurer, J.M., Arbabshirani, M.R., Wilber, C.H., Jaffe, A.J., Sidz, A., Pearlson, G.D., Calhoun, V.D., Clark, V.P., & Kiehl, K.A. (2014) Brain potentials measured during a Go/NoGo task predict completion of substance abuse treatment. *Biological Psychiatry* 76, 75-83.

Stein, M., Fey, W., Koenig, T., Oehy, J., & Moggi, F. (2018) Context-specific inhibition is related to craving in alcohol use disorders: A dangerous imbalance. *Alcoholism: Clinical and Experimental Research* 42, 69-80.

Sterne, J.A.C., Sutton, A.J., Ioannidis, J.P.A., Terrin, N., Jones, D.R., Lau, J., Carpenter, J., Rücker, G., Harbord, R.M., Schmid, C.H., Tetzlaff, J., Deeks, J.J., Peters, J., Macaskill, P., Schwarzer, G., Duval, S., Altman, D.G., Moher, D., & Higgins, J.P.T. (2011)

Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *British Medical Journal* 343, d4002.

Tarter, R.E., Kirisci, L., Mezzich, A., Cornelius, J.R., Pajer, K., Vanyukov, M., Gardner, W., Blackson, T., & Clark, D.B. (2003) Neurobehavioral disinhibition in childhood predicts early age at onset of substance use disorder. *American Journal of Psychiatry* 160, 1078-1085.

Taylor, E.M., Murphy, A., Boyapati, V., Ersche, K.D., Flechais, R., Kuchibatla, S., McGonigle, J., Metastasio, A., Nestor, L., Orban, C., Passetti, F., Paterson, L., Smith, D., Suckling, J., Tait, R., Lingford-Hughes, A.R., Robbins, T.W., Nutt, D.J., Deakin, J.F., & Elliott, R. (2016) Impulsivity in abstinent alcohol and polydrug dependence: a multidimensional approach. *Psychopharmacology (Berl)* 233, 1487-99.

Townshend, J.M., & Duka, T. (2005) Binge drinking, cognitive performance and mood in a population of young social drinkers. *Alcoholism: Clinical and Experimental Research* 29, 317-325.

Trommer, B.L., Hoeppner, J.-A.B., Lorber, R., & Armstrong, K.J. (1988) The Go-NoGo paradigm in Attention Deficit Disorder. *Annals of Neurology* 24, 610-614.

United States Food and Drug Administration (USFDA) (1993) Guideline for the study and evaluation of gender differences in the clinical evaluation of drugs; notice. Federal Register, 58, 39406-39416.

van der Plas, E.A.A., Crone, E.A., van den Wildenberg, W.P.M., Tranel, D., & Bechara, A. (2009) Executive control deficits in substance-dependent individuals: A comparison of alcohol, cocaine, and methamphetamine and of men and women. *Journal of Clinical and Experimental Neuropsychology* 31, 706 - 719.

van Duijvenbode, N., Didden, R., Korzilius, H.P.L.M., Trentelman, M., & Engels, R.C.M.E. (2013) Executive control in long-term abstinent alcoholics with mild to borderline intellectual disability: The relationship with IQ and severity of alcohol use-related problems. *Research in Developmental Disabilities* 34, 3583-3595.

Van Etten, M.L., Neumark, Y.D., & Anthony, J.C. (1999) Male-female differences in the earliest stages of drug involvement. *Addiction* 94, 1413-1419.

Watson, T.D., Newton-Mora, M., & Pirkle, J. (2016) Event-related potential correlates of processing alcohol-related pictures in young adult binge drinkers. *The American Journal of Drug and Alcohol Abuse* 42, 77-87.

Weafer, J., De Arcangelis, J., & de Wit, H. (2015) Sex differences in behavioral impulsivity in at-risk and non-risk drinkers. *Frontiers in Psychiatry* 6, Article 72.

Weafer, J., & de Wit, H. (2014) Sex differences in impulsive action and impulsive choice. *Addictive Behaviors* 39, 1573-1579.

Wessel, J.R. (2018) Prepotent motor activity and inhibitory control demands in different variants of the go/no-go paradigm. Psychophysiology, 55, e12871.

Wetherington, C.L. (2007) Sex-gender differences in drug abuse: A shift in the burden of proof? *Experimental and Clinical Psychopharmacology* 15, 411-417.

White, A., Castle, I.J.P., Chen, C.M., Shirley, M., Roach, D., & Hingson, R. (2015) Converging patterns of alcohol use and related outcomes among females and males in the United States, 2002 to 2012. *Alcoholism: Clinical and Experimental Research* 39, 1712-1726.

Wizemann, T.M., & Pardue, M.-L. (2001) Exploring the biological contributions to human health: Does sex matter? Washington DC, National Academy Press.

World Health Organization (WHO) (2011) Global status report on alcohol and health. Geneva, World Health Organization.

World Health Organization (WHO) (2014) Global status report on alcohol and health. Geneva, World Health Organization.

Wright, L., Lipszyc, J., Dupuis, A., Thayapararajah, S.W., & Schachar, R. (2014) Response inhibition and psychopathology: A meta-analysis of Go/No-Go task performance. *Journal of Abnormal Psychology* 123, 429-439.

Zamorano, F., Billeke, P., Hurtado, J.M., Lopez, V., Carrasco, X., Ossandon, T., & Aboitiz, F. (2014) Temporal constraints of behavioral inhibition: relevance of inter-stimulus interval in a Go-Nogo task. *PLoS One* 9, e87232.

6. ACKNOWLEDGEMENTS

This study was funded by an Australian Rotary Health Postdoctoral Research Fellowship to Dr Smith. The National Drug and Alcohol Research Centre at the University of NSW is supported by funding from the Australian Government under the Substance Misuse Prevention and Service Improvements Grants Fund.

We thank study authors who provided us on request with data suitable for meta-analysis, as well as anonymous reviewers who provided comments on an earlier draft of this manuscript. We also thank Mrs Mary Kumvaj, the NDARC Librarian who assisted with the literature search.