RUNNING HEAD: Working Memory Training and Executive Function

The Effects of Working Memory versus Adaptive Control Training on Executive Cognitive Function

Peter R. Finn, Ph.D.

Luca Nemes, B.A.

Allen Bailey, Ph.D.

Rachel L. Gunn, Ph.D.

Elizabeth A. Wiemers, Ph.D.

Thomas S. Redick, Ph.D.

Peter R Finn (https://orcid.org/0000-0002-0998-0442), Luca Nemes, and Allen Bailey (https://orcid.org/0000-0003-2844-495X), Department of Psychological and Brain Sciences, Indiana University Bloomington; Rachel L. Gunn (https://orcid.org/0000-0001-5244-1105), Center for Alcohol and Addiction Studies, Brown University School of Public Health; Elizabeth A. Wiemers (https://orcid.org/0000-0003-4885-6281), University of Texas at Arlington, and Thomas S. Redick, Department of Psychological Sciences, Purdue University. Correspondences concerning this article should be addressed to Peter Finn at 1101 East 10th Street, Bloomington, IN 47405, USA. finnp@indiana.edu

Acknowledgements: This research was supported by the National Institute on Alcohol Abuse & Alcoholism, grant R01AA13650 to Peter R. Finn (Subcontract: Redick). Manuscript preparation was also supported in part by T32 MH103203 (Bailey), T32 AA07462 (Gunn), K08 AA027551 (Gunn), and NSF Award #1632403 (Redick and Wiemers).

Data availability. The data used in this report is available upon request to Peter R. Finn, finnp@indiana.edu, 1101 East 10th Street, Bloomington, IN 47405, USA

Abstract

There is considerable debate about whether Working Memory (WM) training specifically results in far-transfer improvements in executive cognitive function (ECF) rather than improvements on tasks similar to the training tasks. There has also been recent interest in whether WM training can improve ECF in clinical populations with clear deficits in ECFs. The current study examined the effects of WM training compared with non-WM adaptive Visual Search (VS) control training (15 sessions over 4 weeks) on various measures of ECF, including delay discounting (DD) rate, inhibition on flanker, color and spatial Stroop tasks, and drinking in a community-recruited sample with Alcohol Use Disorder (AUD, 41 men, 41 women, mean age = 21.7 years), who were not in treatment or seeking treatment, and non-AUD healthy controls (37 men, 52 women, mean age = 22.3 years). Both WM and VS training were associated with improvements on all ECF measures at 4 weeks and 1-month follow-up. WM and VS training were associated with reductions in both DD rates and interference on Stroop and Flanker tasks in all participants, as well as reductions in drinking in AUD participants that remained apparent one month post training. The results suggest that nonspecific effects of demanding cognitive training, as opposed to specific WM training effects, could enhance ECF, and that such enhancements are retained at least one-month post-training.

Keywords: working memory training, executive cognitive function, delay discounting, inhibitory control, alcohol use

There is an extensive literature on the impact of working memory (WM) training on various measures of executive cognitive function (ECF), fluid intelligence, and other far transfer cognitive outcome measures (Au et al., 2015), including studies with clinical populations, such as those with Substance Use Disorders (SUDs) or Attention-Deficit-Hyperactivity Disorder (ADHD), exploring the possibility that WM training may affect relevant clinical outcomes (Bickel, Yi, Landes, Hill, & Baxter, 2011; Hendershot et al., 2018; Holmes et al., 2010; Klingberg et al., 2005). A number of WM training studies with SUD samples have examined whether training is associated with reductions in impulsive decision-making, improvements in overall ECF (e.g., Bickel et al., 2011; Hendershot et al., 2018), improvements on measures of self-regulaton (Brooks et al., 2017), or changes in specific neural circuits (Brooks et al., 2016). Overall, the results are mixed. Some studies suggest that WM training is associated with improvements on self-reported ECF measures of self-regulation (e.g., impulsivity: Brooks et al., 2017), while other studies suggest that training does not result in changes in far transfer cognitive measures of ECF (Hendershot et al., 2018; Rass et al., 2015; Snider et al 2018). There is also considerable debate regarding whether WM training improves latent WM capacity and far transfer measures of fluid intelligence (Au, Buschkuehl, Duncan, & Jaeggi, 2016), or simply leads to improvements on tasks that are similar to training tasks (Melby-Lervåg, Redick, & Hulme, 2016).

Executive WM is considered to be a core process underlying ECF in general (Kimberg & Farah, 1993), and is integral to self-control and decision-making. Executive WM includes the attentional control processes (Barrett, Tugade & Engle, 2004; Cowan, 1995; Redick et al., 2016) involved in deliberation during decision-making (Endres, Donkin, & Finn, 2014; Finn, Gunn, & Gerst, 2015) and other executive functions, such as response inhibition (Redick et al., 2016). The

central assumption underlying the current study is that increasing WM capacity via WM training will result in improvements in attention control, such as decreased distractibility and increased capacity to keep in mind less salient information, and the executive functions that rely on attention control / WM capacity, such as decision-making and inhibitory control. Furthermore, because SUD is associated with lower levels of ECF (Finn et al., 2009), increasing WM capacity via WM training might be a useful intervention in populations that have clinically relevant deficits in ECF, such as those with an SUD (Finn et al., 2009). Higher levels of executive WM capacity are associated with superior executive function, including lower levels of impulsive decision-making (i.e., lower delay discounting [DD] rates) (Bobova et al., 2009; Finn et al., 2015; Shamosh et al., 2008), better performance on tasks that assess inhibitory control, such as Stroop (Kane & Engle, 2003; Redick et al., 2016; Unsworth & McMillan, 2014), Flanker tasks (Redick et al., 2016), Go/No-Go tasks (Wiemers & Redick 2019), and better approach-avoidance learning (Endres, Rickert, Bogg, Lucas, & Finn, 2011). Executive WM is typically assessed with *complex dual span tasks*, such as the Operation Span task (Engle et al., 1999), which requires the retention of increasingly larger sets of information while simultaneously engaging in a separate problem-solving task involving encoding and retrieval (Finn et al., 2015; Gunn et al., 2018; Harrison et al., 2013; Wiemers & Redick, 2019).

However, the results of studies of the effects of WM training on various measures of cognitive capacity are mixed. Overall, studies in clinical and nonclinical samples suggest that WM training does not increase latent WM capacity. Rather, studies suggest that WM training leads to specific improvements on WM transfer tasks similar to the training tasks (Gunn et al., 2018; Hendershot et al., 2018; Melby-Lervåg, Redick, & Hulme, 2016; Rass et al., 2015; Snider et al., 2018), but not on transfer tasks that are different in form from the training tasks (Gunn et

al., 2018; Rass et al., 2015; Wanmaker et al., 2018). In fact, some studies report that both active WM and control training are associated with improvements on select transfer tasks (Hendershot et al., 2018; Rass et al., 2015; Wanmaker et al., 2018), while others show no effect of training on WM transfer/cognitive outcome measures (Bickel et al., 2011; Brooks et al., 2017).

Studies also do not provide consistent support for the idea that WM training actually improves decision-making. Of the training studies which assessed DD outcomes, one reported that WM training reduced DD rates (Bickel et al., 2011), while three studies found that training had no effect on DD rates (Hendershot et al., 2018; Rass et al., 2015; Snider et al., 2018). However, one study reported that training reduced DD rates on an episodic future thinking cue facilitated DD task in those initially high in episodic future thinking DD (Snider et al., 2018). Additionally, studies suggest that WM training does not lead to improvements on other decision tasks in SUDs, such as the Balloon Analogue Risk Task (Bickel et al., 2011), the Iowa Gambling Task (Rass et al., 2015), or on measures of response inhibition on Go/No-Go tasks (Bickel et al., 2011; Rass et al., 2015) and Stroop tasks (Schulte et al., 2018). However, WM training has been associated with improvements on self-report measures of self-regulation, impulsivity, and planning in methamphetamine addicts (Brooks et al., 2017), suggesting that training may lead to improvements in the cognitive control components of ECF. There also is some evidence that WM training may increase basal ganglia function (Brooks et al., 2016) and reduce activation in frontal-parietal and striatal networks (Brooks, MacKenzie-Phalen, Tully, & Schioth, 2020), but these effects have not been linked to improved ECF. Finally, there is also some evidence of WM training-related reductions in substance use in outpatient or community samples with a SUD (Houben, Wiers, & Jansen, 2011; Rass et al., 2015), but not in inpatient SUD samples (Hendershot et al., 2018; Wanmaker et al., 2018).

Because WM training studies have suffered from a number of design weaknesses, firm conclusions about the impact of WM training on executive functions in SUDs and other populations cannot be drawn. For instance, the majority of WM training studies do not incorporate an adequate comparison control training (e.g., Beck et al., 2010; Bickel et al., 2011; Brooks et al., 2017; Hendershot et al., 2018; Holmes et al., 2010; Rass et al., 2015; Snider et al., 2018; Wanmaker et al., 2018). With the exception of Gunn et al. (2018), studies of WM training in persons with a SUD have compared adaptive WM training to an undemanding, 2-span training condition that does not involve the same degree of training, comparable effort, or the possibility of improvement. WM training may result in nonspecific improvements in cognitve capacity, because engaging in cognitive training exercises that require substantial effort and involve progressive improvements over an extended period of time leads to improvements across specific task-related skills and effort allocation and heightened sense of mastery, all of which may result in improvements in performance on a range of cognitive tasks. The inclusion of a demanding adaptive cognitive training condition can provide some control for these features of a WM training protocol.

The current study is designed to directly address these limitations by comparing the effects of WM training to a demanding, adaptive, non-WM training (Visual Search—VS) control training condition that has shown improvement with practice without significant transfer to other WM measures (e.g., Redick et al., 2013). Melby-Lervåg and colleagues (2016) conclude that having an adaptive cognitive (non-working memory) training condition is important to control for the amount of required effort, the experience of success and mastery in progresssing to more demanding performance requirements, and any expectancies (i.e., placebo) for the overall impact of training. Indeed, the presence or absence of far transfer after WM training in the meta-analysis

by Melby-Lervåg et al. (2016) depended on whether an adaptive or passive control comparison was used. However, it is important to point out that others (e.g., Au et al., 2020) *do not agree with the necessity for a demanding, adaptive control training comparison condition, or what constitutes the optimal control condition.*

Other design issues in WM training studies include inconsistencies in training sessions (Bickel et al., 2011; Hendershot et al., 2018), very few additional transfer measures (Snider et al., 2018), and a failure to assess whether any observed WM training transfer effects persist for a significant time period post training. If WM training is specifically improving latent WM capacity, then WM training, and *not* control training, should result in significant improvements on different transfer measures immediately post-training, and these effects should persist at post-training follow-up.

The current study was designed to address the limitations in study design noted above and investigate the effect of WM training on a range of ECF measures. A preliminary report of this study appears in Gunn et al. (2018) where we reported on the effects of WM and VS control training on near and moderate transfer measures of WM capacity with a somewhat smaller sub-sample. We did not report the data on the effects of WM training on far transfer measures of ECF at that time, because the study was still ongoing. Near transfer WM measures were tasks that assess WM that were very similar in structure to the tasks involved in WM training, while moderate transfer measures were WM tasks that differed from the WM training tasks in task structure (Gunn et al., 2018). Far transfer measures, such as delay discounting and stroop/flanker interfererence, assessed different executive functions significantly influenced by WM, but are not specific measures of WM capacity. The current paper is the final report of the Gunn et al. (2018), with the addition of 26 subjects, and reports on the effects of WM and adaptive VS

control training on various far transfer measures of ECF. In Gunn et al. (2018), we found that both AUDs and healthy controls showed significant and substantial improvements on WM and VS control training over the 15 training sessions. Gunn et al. (2018) found that WM training, but *not* VS control training, was associated with improvements on near transfer measures of WM capacity, but there was no evidence that WM training was associated with greater improvements on moderate transfer WM measures compared with VS control training. Near transfer WM measures were tasks that assess WM that were very similar in structure to the tasks involved in WM training, while moderate transfer measures were WM tasks that differed from the WM training tasks in task structure (Gunn et al., 2018). The Supplemental Materials Section 1 includes a reanalysis of the Gunn et al. (2018) analyses on the effects of WM and VS control training on near and moderate transfer measures with the current larger sample. The reanalysis of the WM results initially reported in Gunn et al. (2018) did not yield different results from what was originally reported.

The overarching aim of the current study is to test the hypothesis that training on demanding adaptive complex-span WM tasks specifically enhances ECF function compared with VS training, both immediately post training and at 1 month follow-up, in individuals with an AUD as well as healthy controls. In order to test this hypothesis, we investigated the effects of WM training on far transfer ECF measures, such as decision-making (delay discounting rate), inhibitory control (Stroop and Flanker tasks), and alcohol consumption in a large sample of young adult participants with an Alcohol Use Disorder (AUD, n = 82), who are not in treatment or seeking treatment, and a sample of healthy controls (n=89) without AUD. This study was not designed as a clinical trial. Our community-recruited sample was comprised of young adults who are not seeking treatment or cognitive training. The analyses reported here extend the literature

on the effects of WM training in SUD populations and address the pressing question in the broader WM training literature about whether WM training specifically enhances ECF. The analyses were designed to test the hypotheses that adaptive WM training, but not adaptive VS control training, will enhance ECF functioning immediately post training and at one month follow-up evidenced by reductions in DD rates and in interference on Stroop (verbal and spatial) and Flanker tasks across both AUD and non-AUD subjects. We assessed the effects of training on alcohol use as an exploratory/secondary analysis, because other studies of WM training in samples with a substance use disorder assessed training effects on substance use (Hendershot et al., 2018; Rass et al., 2015).

Commented [WE1]: As stated in the reviewer reply comments, I don't think this should've been changed. This seems dishonest to say it was exploratory if it really was a hypothesis. The reviewer was confused.

Commented [RTS2R1]: See my reply to the comment about this in the response to Reviews – but in short, the alcohol self-report was never intended to be a real target of the training.

Method

Participants

Recruitment. Subjects were recruited from the community and campus in Bloomington, Indiana using advertisements and flyers that Finn and colleagues have used in the past (Finn et al., 2015; 2017) that look for "impulsive individuals", "heavy drinkers", "individuals interested in psychological research", and "quiet, introspective, and reflective individuals" (Gunn et al., 2018). The advertisements and flyers did not describe the study as a cognitive training study or look for people seeking treatment or aiming to improve their cognitive abilities, since placebo effects in cognitive training are known to be associated with advertisements that tout the benefits of cognitive training (Foroughi et al., 2016). Individuals who called in response to advertisements were administered a telephone interview to initially screen respondents for the inclusion criteria (noted below) before being invited for a detailed interview process. During the telephone screening, potential participants were told that they were being recruited for "a

research study of personality, emotion, and cognitive factors and alcohol and drug use in young men and women."

Inclusion criteria. To participate in the study, individuals had to (a) be 18 to 30 years of age, (b) read/speak English, (c) have at least a 6th grade education, (d) meet the diagnostic criteria for either the AUD or healthy control groups (noted below), (e) not in treatment or seeking treatment for a SUD, (f) have no history of psychosis or head trauma, (g) available to attend 19 testing sessions, which included 15 training sessions (4 times per week for 4 weeks), 2 baseline assessment sessions, and 2 outcome assessment sessions (one immediately post training and one 4 weeks post training), and (h) agree to complete a urine screen and Breathalyzer test at the beginning of each session. Urine drug screens were used to corroborate self-report data and confirm group inclusion criteria.

Sample selection. Individuals meeting the study inclusion criteria after the telephone screening interview were scheduled for the baseline testing sessions to assess group diagnostic criteria and the other study inclusion criteria. The Semi-Structured Assessment for the Genetics of Alcoholism–version IV (SSAGA-IV: (COGA, 2005) was used to assess the diagnostic group inclusion criteria for the AUD (current AUD) and healthy control group (no history of AUD, other SUD, Conduct Disorder, or Attention-Deficit Hyperactivity Disorder) and lifetime symptom counts (the total number of positive responses to SSAGA questions for each diagnostic category). A total of 221 individuals met the full inclusion criteria for the study and were randomized by group into the WM training and adaptive VS control training conditions. Of the 221 participants, 112 were randomized into WM training (63 AUD; 49 healthy control), with 89 (40 AUD, 49 healthy control) completing the study. 109 participants were randomized into the VS control training condition (62 AUD; 47 healthy control), with 82 (42 AUD, 40 heathy

control) completing the study. Participants were blind to their training condition, but the experimenters were not. There were significantly more dropouts in the AUD group (34.4%) compared with the healthy control group (7.3%), χ^2 (1) = 22.8, p = 0.0001, but there were no significant differences in dropouts between training conditions, χ^2 (1) = 0.57, p = 0.45. Participants dropped out because they were unable to attend the scheduled laboratory sessions on a regular basis. Participants in the AUD group that dropped out of the study did not differ in age (M = 21.6, SD = 2.4 versus M = 21.7, SD = 2.7, t(123) = -.19, p = .85) or lifetime alcohol problems (M = 41.5, SD = 18.3 versus M = 41.0, SD = 18.6, t(123) = .15, p = .89), but they had modestly fewer years of education (M = 14.1, SD = 1.1 versus M = 14.5, SD = 1.0; t(123) = -2.10, p = .038) compared with those that completed the study.

Sample characteristics. The sample consisted of 82 individuals with AUD (41 women), who were not in treatment or seeking treatment, with 39 completing WM training and 43 completing the VS control training. The healthy control group had 89 individuals (52 women) with 49 completing WM training and 40 completing VS control training. Table 1 lists the demographic characteristics, drinking habits (average weekly frequency of drinking and average quantity consumed per drinking day in a typical week over the 3 months prior to the study), and diagnostic lifetime SUD problem counts for each group by each training condition. There were no significant group or training condition differences in age (p=.94), years of education (p=.12), or gender composition (p=.27). Overall, 84.8% of participants were current college students. Lifetime problem counts were derived from SSAGA interview as the total number of positive responses to interview questions in the AUD, Cannabis Use Disorder and Other Drug Use Disorder sections. The racial composition of the sample was 78.1 % Caucasian, 10.6% Asian, 6.3% African-American/Black, and 5% biracial or other. This study was reviewed and approved

by the Indiana University–Bloomington Institutional Review Board (IRB: protocol #0709000094).

INSERT TABLE 1 ABOUT HERE

Procedures

Participants were told that they were doing the training tasks in order to improve performance and determine how much they could improve their performance on the tasks, and it was important to put effort into the tasks. Participants visited the laboratory for a total of 19 sessions: 2 baseline testing sessions, 15 training sessions, and 2 outcome assessment sessions (1 immediately after training, and one at 4 weeks post-training). The two baseline test sessions were used to determine full inclusion and group eligibility, current alcohol use, and assess baseline measures in the three outcome domain areas: 1. delay discounting (DD), 2. cognitive inhibition / interference (two Stroop tasks and a Flanker task), and 3. alcohol consumption (assessed over the previous 2-week period). Following the completion of baseline sessions, participants were scheduled for their training sessions. Each participant completed a total of 15 training sessions at a rate of 4 sessions per week. All sessions were completed in the laboratory. The outcome assessment sessions included the assessments on all outcome measures (see below) at 4 weeks (immediately post training) and 8 weeks (30 days post-training). Participants were monetarily compensated to increase the incentive to attend all sessions and put maximal effort into the training tasks, consistent with other training research (e.g., Redick, Wiemers, & Engle, 2020). Time in the laboratory was compensated at $12/hour and bonuses were paid for arriving on time for each session ($12/session) and for completing the entire study ($20). There were monetary rewards for each time the subject moved to a higher level of difficulty on the training tasks (up to $21 total). If subjects did not complete the study, their hourly pay was reduced to $10/hour and

Commented [WE3]: (See below) as in the response to reviewers is unnecessary. This is now abundantly clear when followed up immediately with the compensation schedule. Further, a "substantial" amount raises alarm bells in that per IRB/APA ethical guidlines, we are not meant to compensate subjects to such a high degree that they feel they 'must' participate, seems better to avoid this language.

on-time arrival bonuses were reduced to $5/session. Subjects could earn up to $830 for completing all sessions and tasks.

Adaptive Training Tasks

Both WM and VS control training tasks were adaptive such that the task difficulty (i.e., set size) increased to a higher difficulty level with successful performance at the current set size. Participants were told the goal was to improve on the tasks. Monetary incentives were provided for improvements from one level to the next in both training protocols. Substantial improvements for all participants were observed on all training tasks (cf., Gunn et al., 2018). Table 1 lists the average score at the final session (training session 15) and the highest level achieved in training by group and training condition.

WM training. WM training consisted of adaptive versions of Operation Span (OS) and Symmetry Span (SS) tasks (Harrison et al., 2013)**.** In the OS task, participants were presented with a series of trials in which they were asked to make an accuracy judgment on an arithmetic equation (e.g., *(6/2) + 1 =5, YES or NO")* before being presented with a to-be-remembered letter (consonant). The task was presented in an adaptive manner, where memory set size (number of equations + letter strings) was increased as the subject improved on the task. All training began at Level 1 difficulty (set sizes 2-4), and set size increased by 1 if the subject was successful on 87.5% or more of the equations and letters across 3 trials (each training session is composed of 8 sets of 3 trials). If a subject correctly recalled 75% or less of the letters, the level of the next set dropped by one in each training session, and the level began at the highest level achieved in the previous session. In the adaptive SS task, participants made symmetry judgments on matrix patterns and recalled matrix locations in the correct serial order. In each trial, participants were presented with a pattern of black and white squares and asked to make a judgment of symmetry

from the vertical axis. Then, participants were presented with a 4x4 matrix with one square highlighted in red, which they were instructed to remember. The number of judgments made and matrix positions per level were the same as the OS task, and the task had identical criteria for level progression.

VS Control training. The VS control training tasks were adaptive visual search (VS) tasks (hands and letters). As evidenced by other research (Foster et al., 2017; Harrison et al., 2013; Redick et al., 2013), these tasks result in training-related gains, without significant transfer to measures of individual WM capacity. In previous studies, our VS training protocol has not resulted in improvements on any transfer measures of WM capacity (Foster et al., 2017; Harrison et al., 2013; Redick et al., 2013; Redick et al., 2020) or far transfer measures of intelligence (Foster et al., 2017; Harrison et al., 2013; Redick et al., 2013). Note also that other research groups (e.g., De Simoni & von Bastian, 2018) have used adaptive VS training as a comparison in WM training studies. Each trial began with a fixation dot in the center of the screen, followed by an array of letters/hands for 500ms, then a mask (16x16 array of black squares) for 2500ms. The size of the letter/hand array varied based on difficulty, beginning with a 2x2 array and advancing based on performance to a maximum of 16x16. Participants were instructed to indicate which direction the target stimulus was facing during the mask presentation. Each array size was presented for 24 trials (1 block), and each training session involved 16 blocks. As in the training tasks, trials increased in difficulty when the subject was accurate on 87.5% of the trials and decreased in difficulty if the subject had less than 75% accuracy.

Outcome Measures

All outcome measures were assessed at baseline (pre-training), 4 weeks (immediately post-training) and at 1-month follow-up (8 weeks).

Delay discounting. The DD task was delivered via desktop computer. Participants were presented with a choice between a specific amount of money "TODAY" or $50 "LATER" at one of six time delays (i.e. 1 week, 2 weeks, 1 month, 3 months, 6 months, 1 year). The immediate choice amount varied from $5.00 to $45.00 in $5.00 increments. Prior to starting the tasks, participants were informed that all money was hypothetical, but were instructed to choose as if they would receive their chosen value in the corresponding time delay. For this task, participants completed 6 blocks, one for each time delay (1 week, 2 weeks, 1 month, 3 months, 6 months, 1 year). Within each block, there were ascending and descending value trials (both the order of the blocks and order of trial type was randomized). In the ascending trials, the immediate reward value began at $5.00 and increased to a maximum of $45.00 in increments of $5.00. The ascending sequence of trials stopped when a participant switched from the delayed to the immediate reward value (or stopped at $5.00 if the immediate reward was chosen on the first trial). There were a total of 9 possible ascending trials for each of the 6 time delay lengths. The point at which participants switched from the delayed value ($50.00) to the immediate option was recorded as the switch point on the ascending trials. On the descending trials, the immediate reward value began at $45.00 and decreased to a minimum of $5.00 in increments of $5.00. For the descending sequence trials, the task stopped when the participants switched from the immediate reward value to the delayed option. The point at which they switched from the immediate to the delayed option ($50) was recorded as their switch point for the descending sequence of trials. Consistent with the ascending trials, there were a maximum of 9 possible trials in the descending sequence for each of the 6 time delay lengths.

As in our previous studies (Bailey, Gerst, & Finn, 2018; Finn et al., 2015; Gerst, Gunn, & Finn, 2017), a single-parameter hyperbolic function was used to estimate discounting rate in both

reward and loss tasks (Mazur, 1987). The estimation of discounting rate was calculated using the following equation: $Vp = V/(1 + k \times dt)$, where Vp is the present (discounted/subjective) value (calculated as the average of the switch points for ascending and descending trials at a particular time delay), the constant V is the amount of the delayed reward ($50), dt is the length of the time the reward or loss is delayed in days, and k is the discounting rate. The estimated k values of each participant was $\log_{10}$ transformed, and this transformed k was used in the subsequent analyses. Consistent with our previous studies, DD data were not excluded for increased variability or extremes in discounting because variation in choice switch points reflects normal variation in DD decisions, and always discounting the delayed reward reflects a legitimate choice (Bailey et al., 2018; Finn et al., 2015; Gerst et al., 2017).

Inhibitory control tasks. Interference measures on two different Stroop tasks and a Flanker task were included as measures of the inhibitory control executive control processes. These tasks assess the attentional inhibitory processes in inhibitory control (Howard, Johnson, & Pascual-Leone, 2014). Interference reflects the degree to which reaction times (RTs) are slower on incongruent trials relative to neutral trials. Smaller interference RTs are thought to reflect a superior ability to inhibit, or suppress, the task irrelevant processing requirement / dominant response sets features of incongruent trials (Friedman & Miyake, 2004; Howard et al., 2014).

The color - word Stroop task required participants to identify the color that a word is presented in (red, green, or blue), while ignoring the identity of the word. The color Stroop interference effect refers to the typical observation that participants are slower to respond on incongruent trials (e.g., the word ‘red’ presented in green font) versus neutral trials (a non-color word, such as tree, is presented in a color) and congruent trials (e.g., the word ‘red’ presented in red font). The task began with 15 response-mapping trials, with the answers ‘red’, ‘green’, and

'blue' mapped to the 1, 2, and 3 keys on the number keypad on the keyboard. The number keys were covered with stickers representing their respective colors. Then, subjects performed 6 practice trials (3 congruent and 3 incongruent trials). The experimental block consisted of 144 trials (108 congruent/18 incongruent/18 neutral). Congruent trials made up 75% of the overall number of trials because previous research (e.g., Kane & Engle, 2003) has shown that color Stroop interference effects are stronger when the color and word information are frequently consistent. Colors and words were presented equally often within each trial type. On each trial, words were presented until participants responded, followed by a blank screen presented between trials for 1000, 1500, or 2000 ms. Accuracy feedback was presented on practice trials only. For the analyses, interference RT was quantified as the RT on incongruent trials minus the RT on neutral trials.

The spatial Stroop task required participants to respond to the direction that an arrow is pointing (left or right), while ignoring the arrow's onscreen location (left, center, or right). The spatial Stroop effect refers to the typical observation that subjects are slower to respond on incongruent trials (e.g., a left arrow on the right side of the screen) versus congruent trials (e.g., a right arrow on the right side of the screen) and neutral trials (the arrow presented centrally). The task procedure was similar to the one used in Redick et al. (2016). The task began with 10 response-mapping neutral trials. Then, participants performed 4 practice trials (2 congruent and 2 incongruent trials). The experimental block was 144 trials (108 congruent/18 incongruent/18 neutral). Congruent trials made up 75% of the overall number of trials because previous research (e.g., Logan & Zbrodoff, 1979) has shown that spatial Stroop interference effects are stronger when arrow direction and location are frequently consistent. Left and right arrows were presented equally often within each trial type. On each trial, arrows were presented for up to

3000 ms, and then a blank screen appeared between trials for 1500, 2000, or 2500 ms. Accuracy feedback was presented on practice trials only. Participants made responses by pressing the "z" and "/" keys, which were labeled with left and right arrow stickers, respectively. For the analyses, interference RT was quantified as the RT on incongruent trials minus the RT on neutral trials.

The arrow Flanker task required participants to respond to the direction that a central arrow is pointing (left or right) while ignoring the arrows that flank the center target on either side. The flanker effect refers to the typical observation that subjects are slower to respond on incongruent trials (e.g., a left arrow surrounded by right arrows) versus congruent trials (e.g., a right arrow surrounded by right arrows) or neutral trials (e.g., centrally presented arrow and no arrowheads presented on flanking stimuli). The task procedure was similar to the one used in Redick et al. (2016). The task began with 10 response-mapping neutral trials. Then, participants performed 6 practice trials (2 congruent, 2 incongruent, and 2 neutral trials). The experimental block was 150 trials total (50 congruent/50 incongruent/50 neutral). Left and right arrows were presented equally often within each trial type. A fixation (+) was present in the center of the screen throughout each trial. Each trial began with a 400 ms fixation display, followed by an asterisk cue above the fixation for 100 ms, and then another fixation display for 400 ms. The five arrows (1 target and 4 flanker) were then shown until the subject responded (maximum of 1700 ms). There was a 400 ms interval before the next trial began. Accuracy feedback was presented on practice trials only. Subjects made responses by pressing the "z" and "/" keys, which were labeled with left and right arrow stickers, respectively. For the analyses, interference RT was quantified as the RT on incongruent trials minus the RT on neutral trials.

Alcohol use outcome measures. The effect of training on alcohol use was assessed with a 2-week time-line follow back procedure (Sobell & Sobell, 1992). Each day over the past 2 weeks was reviewed for drinking. The drinking outcome measures were the mean frequency of drinking occasions (per week) and mean quantity consumed per occasion over the past 2 weeks.

Data Analyses. The data were analyzed using IBM SPSS – 26 (SPSS, 2019) for each outcome domain in two ways. First, ANOVA or MANOVA were run to determine whether there were baseline differences between training conditions within each domain and to confirm that there were no training condition differences in sample characteristics. Second, hypotheses regarding the effects of WM versus VS control training and training condition by group effects (drinking outcomes) across the 3 assessment times were tested with general linear repeated measures (Group by Training Condition by Time) MANOVA or ANOVA (DD rates) within each domain (near WM transfer, moderate WM transfer, DD rates, cognitive inhibition and drinking outcomes). Effects sizes are in Cohen's *d*. When reporting the effects of time (i.e., the effects of training), effect sizes are presented for the full-time main effect from baseline to week 8 (follow-up). Effects at 4 weeks (immediate post-training) are reported in the case of quadratic treatment effects.

Commented [RTS4]: Check this – I think this should be dropped for the revision

Results

Delay Discounting Rates and Training. The analyses (ANOVA: Group by Training Condition) of the baseline DD rates revealed no significant differences in baseline DD rates between groups, $F(1,167) = 0.31$, $p = 0.58$, and no significant differences across training condition, $F(1,167) = 0.98$, $p = 0.32$. The interaction between Group and Training condition was also not significant, $F(1,167) = 1.36$, $p = 0.24$.

In the analysis of the effects of training, the 'Training Condition by Time' interaction was not significant, $F(2,332) = 0.29$, $p = 0.75$. However, the overall main effect of Time was significant, $F(2,332) = 22.1$, $p = 0.0001$. Both WM, $F(2,172) = 11.7$, $p = 0.0001$, $d = -0.37$, and VS control training, $F(2,160) = 10.75$, $p = 0.0001$, $d = -0.34$, resulted in reductions in DD rates (see Figure 1). The overall main effect of Group was also not significant in this analyses, $F(1,166) = 0.98$, $p = 0.34$, indicating that overall, AUD participants did not differ from non-AUD participants in DD rate.

INSERT FIGURE 1 ABOUT HERE

Inhibitory Control and Training. On baseline measures of interference on the Color Stroop, Spatial Stroop, and Flanker tasks (RT incongruent–RT neutral trials), MANOVA (Group by Training Condition with all 3 measures) revealed no Group differences, $F(3,164) = 2.02$, $p = 0.113$, no Training condition differences, $F(3,164) = 0.46$, $p = 0.71$, and no significant Group by Training condition interaction, $F(3,164) = 0.98$, $p = 0.404$, on measures of inhibitory control.

The analysis of the effects of training revealed no significant multivariate Training Condition by Time interaction, $F(6,161) = 1.27$, $p = 0.27$, while the multivariate effect of Time was significant, $F(6,161) = 8.25$, $p = 0.0001$, indicating that both WM training and VS control training resulted in significant overall reductions in interference on Spatial Stroop, $F(2,332) = 6.26$, $p = 0.002$, $d = -0.31$, Color Stroop $F(2,334) = 10.13$, $p = 0.0001$, $d = -0.45$, and Flanker tasks, $F(2,334) = 19.10$, $p = 0.0001$, $d = -0.43$. Figure 2 presents these results. The analysis indicated no overall multivariate effect of Group, $F(3,164) = 0.61$, $p = 0.612$, suggesting that the AUD participants did not differ from healthy controls on any of the measures of inhibitory control.

INSERT FIGURE 2 ABOUT HERE

The Supplemental Materials Section 2 reports more detailed analyses of RTs on incongruent and neutral trials across Time for each measure. These analyses show reductions in RTs for both types of trials across time in all tasks, but steeper reductions in RTs on incongruent versus neutral trials (p = .001) through week 8 on the Spatial Stroop and Flanker tasks, which is consistent with the finding of decreased interference after training through week 8. The Color Stroop analyses suggest training related reductions in interference from baseline to week 4, but not through to week 8.

Alcohol Consumption and Training. On baseline measures (frequency of drinking per week and average quantity consumed per drinking day over the past 2 weeks), the MANOVA revealed a significant group effect, $F(2,166) = 107.02$, $p = 0.00001$. As expected, AUD participants drank more than healthy controls on both measures. There were no significant effects of Training Condition, $F(2,166) = 2.1$, $p = 0.118$, or Group by Training Condition interaction, $F(2,166) = 1.16$, $p = 0.316$, at baseline.

The MANOVA on the effects of training revealed a significant Group by Time interaction, $F(4,164) = 3.93$, $p = 0.004$. The Time by Training Condition, $F(4,164) = 0.87$, $p = 0.48$, and the Time by Training Condition by Group interactions, $F(4,164) = 0.64$, $p = 0.637$ were not significant. In the AUD group, modest reductions in the frequency of drinking were observed in both the WM training, $F(1,167) = 10.0$, $p = 0.003$, $d = -0.36$, and VS control training conditions, $F(1,167) = 27.3$, $p = 0.0001$, $d = -.63$. In addition, the quantity consumed per occasion was also reduced in both the WM training, $F(1,167) = 4.71$, $p = 0.03$, $d = -0.20$, and VS control training conditions, $F(1,167) = 10.1$, $p = 0.003$, $d = -0.35$. Figure 3 displays these results.

INSERT FIGURE 3 ABOUT HERE

Discussion

The results indicated that both WM training and adaptive VS control were related to improvements on all far transfer measures of ECF, and that these improvements were retained at 1-month follow-up in a sample comprised of approximately 50% who had an AUD diagnosis and 50% who were healthy, non-AUD controls. WM training and VS control training were associated with large and significant reductions in delay discounting (DD) rates, interference on both Stroop tasks and the Flanker tasks, and reductions in excessive drinking in the AUD participants only. Overall, the results suggest that intensive and demanding cognitive training in general may be associated with enhanced performance on a range of far-transfer measures of executive cognitive capacity, but these effects appear to be non-specific and the mechanisms of change remain unclear.

Contrary to the findings of most studies of the impact of WM training on delay discounting (DD), we found significant decreases in DD rates after both WM and VS control training that persisted through to the 1 month follow-up assessment. Further, an unexpected finding was the lack of baseline group differences in DD rates and measures of inhibitory control. Unique to our study is the finding that an adaptive non-working memory control training was also associated with reductions in DD rates, with only one other study to date (i.e., Bickel et al., 2011) reporting that WM training led to decreases in DD rates, while three studies reported no effect of WM training on discounting rates (Hendershot et al., 2018; Rass et al., 2015; Snider et al., 2018). Given that both the WM and VS training protocols were associated with sustained reductions in DD rates, the mechanisms by which training led to general improvements in delay discounting are unclear.

Both WM and VS control training were also associated with reductions in interference on the Flanker and both Stroop tasks. Follow-up analyses using RTs on both incongruent and

neutral trials detailed in the Supplemental Materials confirmed that training led to reductions in interference on all tasks at week 4 and lasted through to week 8 on the Spatial Stroop and Flanker tasks. The literature on the impact of WM training on measures of inhibitory control is not as extensive as those examining the effects of training on various measures of WM and fluid intelligence. Of the studies that have examined the effects of WM training on measures of inhibitory control, the results are mixed. Two studies found that WM training did not affect Stroop performance (Wanmaker et al., 2018; Schulte et al., 2018). Another study found that Flanker task performance improved with WM training (Sari, Kostera, Pourtois, & Derakshan, 2016), and two studies reported that WM training had no impact on Go/No-Go measures of inhibitory control (Bickel et al., 2011; Rass et al., 2015).

Although not a primary focus, the current study also includes exploratory analyses of the effect of WM training on alcohol use. We found that both WM and VS control training were associated with reductions in self-reported drinking in AUD participants. AUD participants drank 1 less drink per occasion (from 8.1 to 7.0 drinks) and drank fewer occasions per week (from 3.4 to 2.6 occasions per week). Although the reductions in drinking were statistically significant, they were not clinically significant as the post-training levels of drinking were still excessive in the AUD participants. AUD participants were still binge drinking on an average of 2.6 occasions per week (7 drinks per occasion). The reductions in drinking in the AUD participants could be associated with the repeated assessments of drinking immediately after training and at follow-up, which might have created the expectancy that training may result in reduced drinking. However, the results say little about the potential value of using WM training as an approach to reduce drinking in persons with AUD given the reductions in drinking were not clinically significant. Critically, we found that VS control training produced similar effects.

Commented [WE5]: alternatively "they were of limited clinical utility"?

Commented [RTS6R5]: I would leave as 'clinically relevant', or change as Liz suggests. But whatever the final choice is, the copy-pasted text in response to reviewer 2 should also be checked to make sure it matches.

Commented [WE7]: again, as stated in the reply letter, I think forming habits or 'guilt' in having to report it might be a more (or at least equally) plausible mechanism for these reductions

Commented [RTS8]: Again, if accept Liz's edits, make sure it matches the text that is copy-pasted in reply to reviewer 2

In fact, among the AUD sample, the reduction in self-reported drinking was numerically larger in the VS control group versus the WM training group (Figure 3). Nonetheless, the reductions in drinking observed in the AUD group are interesting, because participants were not seeking treatment to reduce drinking and were not given the expectation that training would impact their drinking. Thus, it is unclear from our results whether WM training may be useful as an intervention to reduce excessive drinking.

The fact that both the WM and VS control training resulted in improvements in ECF, as well as the moderate transfer WM capacity measures in Gunn et al., 2018 (see section 1 of the Supplemental Materials), suggests that these effects are not specifically due to WM training and may be associated with training practice effects, non-specific training effects, or improvements in cognitive function other than WM. Previous studies indicate that some of the improvements after WM training are associated with practice effects (Redick et al., 2012; Unsworth, Redick, Heitz, Broadway, & Engle, 2009), which may partially explain the pattern of change observed on the outcome tasks after both WM and VS control training protocols. Our results also suggest that some non-specific factors associated with intensive (4 times per week) cognitive training may lead to improvements in performance on cognitive tasks in general; however, the fact that these improvements were retained at 1-month follow-up suggests that the changes could be associated with improvements in some cognitive or functional domains, or perhaps learning new skills, that were long-lasting (Gathercole et al., 2019).

Finally, somewhat unique to the current study are the levels of incentives used to encourage effort on the training tasks which may have contributed to the improved performance on the two training tasks and lead to improved performances on many of the transfer tasks. The stepwise increase in incentives as the participant progressively improved in the training condition

may have increased the effort and attention on the outcome measures. Thus, it is possible that the significant training-related incentives played a role in improved overall performance. Notably, the current study was not a clinical trial-treatment study in which participants enrolled with the desire or expectation of improvement in any area. Participants signed up for a generic study on cognition and were not seeking treatment or training of any kind. Thus, the widespread improvements seen after both training regimens are less likely due to simple expectancy effects.

There are a number of limitations that should be considered when interpreting the results. First, participants were not seeking treatment of any kind and did not enroll in the study with the expectation that training would improve their cognitive abilities. Thus, the question remains open whether WM training has value as a treatment component for individuals actively in treatment or seeking treatment for an AUD or SUD. Second, our sample is largely comprised of college students who were able to attend sessions regularly (4 times per week for the first 4 weeks) for testing in an on-campus laboratory. Third, the study involved monetary incentives, including the incentive to improve on the training tasks. Some have argued that large financial incentives may mitigate against obtaining training-related transfer (Jaeggi et al., 2014); however, this does not appear to be the case in the present study, because both training protocols were associated with significant improvements in ECF. Fourth, one-third of the enrolled AUD participants dropped out of the study.

The final sample of AUD participants may have higher levels of cognitive functioning than is typical for persons with an AUD, which is suggested by the lack of baseline group differences in delay discounting rates, inhibitory control, and WM capacity (see Gunn et al., 2018; Supplemental Materials). Fifth, some research questions the psychometric properties of the Stroop and Flanker tasks as measures of cognitive inhibition. While these tasks consistently

produce reliable and robust differences between congruent and incongruent conditions (Hedge et al., 2018), the use of difference scores to operationally define inhibitory ability can be problematic for individual differences analyses (Draheim et al., 2019). However, whereas select studies find little-to-no overlapping variance among inhibitory measures (e.g., Rey-Mermet et al., 2019), many other studies have identified a strong covariance among inhibitory control measures reflected in a coherent interference control latent variable (e.g., Friedman & Miyake, 2004; Redick et al., 2016; Tsukahara et al., 2020; Unsworth, 2010). Of course, for the current usage as separate transfer tasks in the main analyses, the magnitude of interrelationships among Stroop and Flanker are much less of a concern. Finally, although including a demanding adaptive VS control training comparison condition is important to make firm conclusions about the specific impact of WM training, in retrospect, it may have been useful to include the kind of non-adaptive control comparison condition as an additional training control condition (e.g., simple 2-span tasks repeated over training sessions used in many Cogmed studies). The inclusion of a simple, non-adaptive training condition would have allowed for conclusions about the impact of the demanding nature of both WM and VS control training. In general, the results suggest that nonspecific aspects of WM and VS control training, including the amount of effort involved in the training and the focus on performance, may have led participants to put more effort into the different outcome tasks at weeks 4 and 8.

In summary, the results suggest that WM training does not uniquely impact far transfer measures of ECF. However, the fact that both WM and VS control training were associated with improvements in all measures of ECF that were retained at 1 month post training suggests that the demanding adaptive cognitive training may result in substantive changes in some cognitive or functional domain (such as specific task skills) that are long-lasting. For instance, Brooks and

colleagues (Brooks et al., 2016, 2020) report that WM training results in changes in brain networks that may underlie complex cognitive functions or skills. It is conceivable that in our study, both WM and VS control training conditions led to changes in specific networks (such as frontal-parietal networks) which underlie the post-training changes we observed on measures of ECF. However, our data do not allow for the identification of specific mechanisms by which training may have improved performance on any measure.

Declarations

Funding: This research was supported by National Institutes of Alcohol Abuse & Alcoholism grant R01AA13650 to Peter. R. Finn (Subcontract: Redick). Manuscript preparation was also supported in part by T32 MH103203 (Bailey), T32 AA07462 (Gunn), K08 AA027551 (Gunn), and NSF Award # 1632403 (Redick and Wiemers).

Conflicts of interest/Competing interests. The authors (Peter Finn, Luca Nemes, Allen Bailey, Rachel Gunn, Elizabeth Wiemers, and Thomas Redick) declare that they have no conflict of interest in any aspect of the data collection, analysis, interpretation, or manuscript preparation.

Availability of data and material. The de-identified data used in this report is freely available to any researcher upon reasonable request and for non-commercial purposes. Contact Peter R Finn, PhD at finnp@indiana.edu or at the Department of Psychological and Brain Sciences, Indiana University, 1101 E 10th Street, Bloomington, IN 47405

Code availability; Not applicable

Authors' contributions; Peter Finn designed the study, wrote the manuscript, and conducted the data analysis. Thomas Redick designed the study and wrote sections of the manuscript. Luca Nemes edited and assisted in the writing of the manuscript. Rachel Gunn edited the manuscript

and provided input on the design of the study. Elizabeth Wiemers overall data processing and edited the manuscript.

Ethics approval. This study was reviewed and approved by the Institutional Review Board at Indiana University.

Consent to participate. All participants provided their informed consent to participate in this study.

Consent for publication. The authors consent to publish this report.

References

Au, J., Buschkuehl, M., Duncan, G.J., & Jaeggi, S.M. (2016). There is no convincing evidence that working memory training is NOT effective: a reply to Melby-Lervåg and Hulme (2015). *Psychonomic Bulletin & Review, 23,* 331–337.

Au, J., Gibson, B.C., Bunarjo, K.,Buschkuehl, M., & Jaeggi, S.M. (2020). Quantifying the Difference Between Active and Passive Control Groups in Cognitive Interventions Using Two Meta-analytical Approaches. *Journal of Cognitive Enhancement, 4,* 192-210. https://link.springer.com/article/10.1007/s41465-020-00164-6

Au, J., Sheehan, E., Tsai, N., Duncan, G. J., Buschkuehl,M.,& Jaeggi, S.M. (2015). Improving fluid intelligence with training on working memory: A meta-analysis. *Psychonomic Bulletin and Review, 22*, 366–377. doi:10.3758/s13423-014-0699-x

Bailey, A.J., Gerst, K, & Finn, P.R. (2018) Delay discounting of losses and rewards in Alcohol Use Disorder: The effect of working memory load. Psychology of Addictive Behaviors, 32(2) 197-204.

Barrett, L.F., Tugade, M.M., & Engle, R.W. (2004). Individual differences in working memory capacity and dual-process theories of the mind. *Psychological Bulletin, 130*, 553–573.

Beck, S.J., Hanson, C.A., Puffenberger, S.S., Benninger, K.L., & Benninger, W.B. (2010) A Controlled Trial of Working Memory Training for Children and Adolescents with ADHD. *Journal of Clinical Child & Adolescent Psychology 39*, 825–836.

Bickel, W.K., Yi, R., Landes R.D., Hill, P.F., & Baxter, C. (2011). Remember the future: Working memory training decreases delay discounting among stimulant addicts. *Biological Psychiatry,* 69, 260–265.

Bobova, L., Finn, P. R., Rickert, M. E., & Lucas, J. (2009). Disinhibitory psychopathology and

Commented [RTS9]: De Simoni, C., & von Bastian, C. C. (2018). Working memory updating and binding training: Bayesian evidence supporting the absence of transfer. *Journal of Experimental Psychology: General, 147*(6), 829–858. https://doi.org/10.1037/xge0000453

delay discounting in alcohol dependence: Personality and cognitive correlates. *Experimental and Clinical Psychopharmacology, 17,* 51-61.

Brooks, S.J., Burch, K.H., Maiorana, S.A, Cocolas, E., Schioth, H.B., Nilsson, E.K., Kamaloodien, K., & Stein, D.J. (2016) Psychological intervention with working memory training increases basal ganglia volume: a VBM study of inpatient treatment for methamphetamine use. *Neuroimage Clinical, 12*, 478–491.

Brooks, S.J., Wiemerslage, L., Burch, K., Maiorana, S.A., Cocolas, E., Schiöth, H. B., Kamaloodien, K., Kamaloodien, K., & Stein, D.J. (2017). The impact of cognitive training in substance use disorder: the effect of working memory training on impulse control in methamphetamine users. *Psychopharmacology*, *234*, 1911–1921.

Brooks, S.J., MacKenzie-Phalen, R., Tully, J., & Schioth, H.B. (2020). Review of the neural processes of working memory training: Controlling the impulse to throw the baby out with the bathwater. Front. Psychiatry 11:512761. doi: 10.3389/fpsyt.2020.512761

Brown, J. (1958) Some tests of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology, 10*, 12–21.

Collaborative Study on the Genetics of Alcoholism – COGA (2005). *Semi-Structured Interview for the Genetics of Alcoholism-Version Four (SSAGA-IV)*.

Cowan, N. (1995). *Attention and memory : an integrated framework*. Oxford University Press.

Draheim, C., Mashburn, C. A., Martin, J. D., & Engle, R. W. (2019). Reaction time in differential and developmental research: A review and commentary on the problems and alternatives. *Psychological Bulletin*, *145*(5), 508-535. doi: 10.1037/bul0000192

Endres, M.J., Donkin, C., & Finn, P.R. (2014). An information processing / associative learning account of behavioral disinhibition in externalizing psychopathology. *Experimental Clinical*

Psychopharmacology, *22*, 122-132.

Endres, M. J., Rickert, M. E., Bogg, T., Lucas, J., & Finn, P. R., (2011). Externalizing psychopathology and behavioral disinhibition: Working memory mediates signal discriminability and reinforcement moderates response bias in approach-avoidance learning. *Journal of Abnormal Psychology, 120*, 336-351.

Engle, R.W., Tuholski, S.W., Laughlin, J.E., & Conway, A.R. (1999). Working memory, short-term memory, and general fluid intelligence: a latent variable approach. *Journal of Experimental Psychology: General, 128*, 309-331.

Finn, P.R. (2002). Motivation, working memory, and decision making: a cognitive-motivational theory of personality vulnerability to alcoholism. *Behavioral and cognitive neuroscience reviews, 1*, 183–205.

Finn, P.R., Gunn, R.L. & Gerst, K.R. (2015). The effects of a working memory load on delay discounting in those with externalizing psychopathology. *Clinical Psychological Science, 3*, 202–214.

Finn, P.R., Rickert, Miller, M., M.E., Lucas, J., Bogg, T., Bobova, L., & Cantrell, H (2009). Reduced cognitive capacity in alcohol dependence: Examining the role of covarying externalizing psychopathology. *Journal of Abnormal Psychology, 118*, 100-116.

Foroughi, C. K., Monfort, S. S., Paczynski, M., McKnight, P. E., & Greenwood, P. M. (2016). Placebo effects in cognitive training. *Proceedings of the National Academy of Sciences of the USA, 113* (27), 7470-7474. doi: 10.1073/pnas.1601243113

Foster, J.L., Harrison, T.L., Hicks, K.L., Redick, T.S., & Engle, R.W. (2017). Do the effects of working memory training depend on baseline ability level?. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43,* 1677.

Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control

functions: A latent-variable analysis. *Journal of Experimental Psychology: General*, *133*(1), 101-135. doi: 10.1037/0096-3445.133.1.101

Gathercole, S. E., Dunning, D. L., Holmes, J., & Norris, D. (2019). Working memory training involves learning new skills. *Journal of Memory and Language*, *105*, 19-42. doi: 10.1016/j.jml.2018.10.003

Gerst, K., Gunn, R.L., & Finn, P.R. (2017). Delay discounting of losses in Alcohol Use Disorders and Antisocial Psychopathology: Effects of a working memory load. Alcoholism: Clinical and Experimental Research, 41, 1768-1774

Gunn, R. L., Gerst, K. R., Wiemers, E. A., Redick, T. S., & Finn, P. R. (2018). Predictors of effective working memory training in individuals with Alcohol Use Disorders. *Alcoholism: Clinical and Experimental Research, 42*, 2432-2441.

Harrison, T.L., Shipstead, Z., Hicks, K.L., Hambrick, D.Z., Redick, T.S., & Engle, R.W. (2013). Working Memory Training May Increase Working Memory Capacity but Not Fluid Intelligence. *Psychological Science 24,* 2409–2419.

Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods*, *50*(3), 1166-1186. doi: 10.3758/s13428-017-0935-1

Heinzel, S., Lorenz, R.C., Pelz, P., Heinz, A., Walter, H., Kathmann, N., Rapp, M.A., & Stelzel, C. (2016). Neural correlates of training and transfer effects in working memory in older adults. *Neuroimage*, 134, 236–49. doi: 10.1016/j.neuroimage.2016.03.068

Hendershot, C.S., Wardell, J.D., Vandervoort, J., McPhee, M.D., Keough, M.T., & Quilty, L.C. (2018). Randomized trial of working memory training as an adjunct to inpatient substance use disorder treatment. *Psychology of Addictive Behaviors*, *32,* 861.

Holmes, J., Gathercole, S.E., Place, M., Dunning, D.L., Hilton, K.A., & Elliott, J.G. (2010). Working memory deficits can be overcome: Impacts of training and medication on working memory in children with ADHD. *Applied Cognitive Psychology, 24*, 827–836.

Houben, K., Wiers, R.W., & Jansen, A. (2011). Getting a grip on Drinking Behavior. *Psychological Science, 22,* 968–975.

Howard, S.J., Johnson, J., & Pascual-Leone, J. (2014). Clarifying inhibitory control: diversity and development of attentional inhibition. Cognitive Development, 31, 1-21.

Jaeggi, S.M., Buschkuehl, M., Shah, P., & Jonides, J. (2014). The role of individual differences in cognitive training and transfer. *Memory & Cognition*, *42,* 464–480.

Kane, M.J., & Engle, R.W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General, 132*, 47–70.

Kimberg, D. Y., & Farah, M. J. (1993). A unified account of cognitive impairments following frontal lobe damage: The role of working memory in complex, organized behavior. *Journal of Experimental Psychology: General*, *122*, 411-428.

Klingberg, T., Fernell, E., Olesen, P.J., Johnson, M., Gustafsson, P., Dahlström, K., & Westerberg, H. (2005). Computerized training of working memory in children with ADHD--a randomized, controlled trial. *Journal of the American Academy of child and adolescentant psychiatry, 44,* 177–86.

Logan, G.D., & Zbrodoff, N.J. (1979). When it helps to be misled: Facilitative effects of increasing the frequency of conflicting stimuli in a Stroop-like task. *Memory & cognition, 7*, 166-174.

Melby-Lervåg, M. & Hulme, C. (2013). Is working memory training effective? A meta-analytic

review. *Developmental Psychology*, *49,* 270–291.

Melby-Lervåg, M., Redick, T.S., Hulme, C. (2016). Working memory training does not improve performance on measures of intelligence or other measures of "far transfer" evidence from a meta-analytic review. *Perspectives on Psychological Science*., *11,* 512-534

Rass, O., Schacht, R.L., Buckheit, K., Johnson, M.W., Strain, E.C., & Mintzer, M.Z. (2015). A randomized controlled trial of the effects of working memory training in methadone maintenance patients. *Drug and Alcohol Dependence, 156,* 38-46.

Redick, T.S. (2015). Working memory training and interpreting interactions in intelligence interventions. *Intelligence*, *50,* 14–20.

Redick, T. S., Broadway, J. M., Meier, M. E., Kuriakose, P. S., Unsworth, N., Kane, M.J., & Engle, R. W. (2012). Measuring working memory capacity with automated complex span tasks. *European Journal of Psychological Assessment, 28*, 164-171. doi:10.1027/1015-5759/a000123

Redick T.S., Shipstead, Z., Harrison, T.L., Hicks, K.L., Fried, D.E., Hambrick, D.Z., & Engle, R.W. (2013). No evidence of intelligence improvement after working memory training: A randomized, placebo-controlled study. *Journal of Experimental Psychology: General, 142,* 359–379.

Redick, T.S., Shipstead, Z., Meier, M.E., Montroy, J.J., Hicks, K.L., Unsworth, N., Kane, M.J., Hambrick, D.Z., & Engle, R.W. (2016). Cognitive predictors of a common multitasking ability: Contributions from working memory, attention control, and fluid intelligence. *Journal of Experimental Psychology: General, 145*, 1473–1492.

Redick, T. S., Wiemers, E. A., & Engle, R. W. (2020). The role of proactive interference in

working memory training and transfer. *Psychological Research, 84,* 1635-1654. doi: 10.1007/s00426-019-01172-8

Rey-Mermet, A., Gade, M., Souza, A. S., von Bastian, C. C., & Oberauer, K. (2019). Is executive control related to working memory capacity and fluid intelligence? *Journal of Experimental Psychology: General*, *148*(8), 1335-1372. doi: 10.1037/xge0000593

Sari, B.A., Kostera, E.H.W., Pourtois, G., & Derakshan, N. (2016). Training working memory to improve attentional control in anxiety: A proof-of-principle study using behavioral and electrophysiological measures. Biological Psychology 121 (2016) 203–212

Schulte, M.H.J., Wiers, R.W., Boendermaker, W.J., Goudriaan,A.E., van den Brink, W., van Deursen, D.S., Friese, M., Bredeg, E., & Waters, A.J., (2018). The effect of N-acetylcysteine and working memory training on cocaine use, craving and inhibition in regular cocaine users: correspondence of lab assessments and Ecological Momentary Assessment Addictive Behaviors 79 (2018) 24–31

Shamosh, N.A., DeYoung, C.G., Green, A.E., Reis, D.L., Johnson, M.R., Conway, A.R.A., Engle, R.W., Braver, T.S., & Gray, J.R. (2008). Individual differences in delay discounting: relation to intelligence, working memory, and anterior prefrontal cortex. *Psychological Science, 19*, 904-911.

Shipstead, Z., Redick, T.S., & Engle, R.W. (2012). Is working memory training effective? *Psychological Bulletin, 138,* 628–654.

Snider, S.E., Deshpande, H.U., Lisinski, J.M., Koffarnus, M.N., LaConte, S.M. & Bickel, W.K. (2018). Working memory training improves alcohol users' episodic future thinking: a rate-dependent analysis. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 3,* 160–167.

Sobell, L., & Sobell, M. (1992). Timeline Follow-Back: A technique for assessing self-reported alcohol consumption. In R. Z. Litten & J. Allen (Eds.), Measuring alcohol consumption: Psychosocial and biological methods (pp. 41–72). Totowa, NJ: Humana Press.

Soveri, A., Antfolk, J., Karlsson, L., Salo, B., & Laine, M. (2017). Working memory training revisited: A multi-level meta-analysis of n-back training studies. *Psychonomic Bulletin & Review*, *24*(4), 1077-1096. doi: 10.3758/s13423-016-1217-0

SPSS version 26 (2019). IBM SPSS release 26. IBM Corporation.

Tsukahara, J. S., Harrison, T. L., Draheim, C., Martin, J. D., & Engle, R. W. (2020). Attention control: The missing link between sensory discrimination and intelligence. *Attention, Perception, & Psychophysics*, *82*(7), 3445-3478. doi: 10.3758/s13414-020-02044-9

Unsworth, N. (2010). Interference control, working memory capacity, and cognitive abilities: A latent variable analysis. *Intelligence*, *38*(2), 255-267. doi: 10.1016/j.intell.2009.12.003

Unsworth, N., & McMillan, B.D. (2014). Similarities and differences between mind-wandering and external distraction: A latent variable analysis of lapses of attention and their relation to cognitive abilities. *Acta psychologica, 150*, 14–25.

Unsworth, N., Redick, T. S., Heitz, R. P., Broadway, J. M., & Engle, R. W. (2009).Complex working memory span tasks and higher-order cognition: A latent-variable analysis of the relationship between processing and storage. *Memory, 17*, 635-654. doi:10.1080/09658210902998047

Wanmaker. S., Leijdesdorff, S.M.J., Geraerts, E., van del Wetering, B.J, Renkema, P.J., & Franken, I.H. (2018). The efficacy of a working memory training in substance use patients: A randomized double-blind placebo-controlled clinical trial. *Journal of Clinical and Experimental Neuropsychology*, *40,* 473-486.

Wiemers, E.A., & Redick, T.S. (2019). Task manipulation effects on the relationship between working memory and go/no-go task performance. *Consciousness and Cognition, 71,* 39-58.

Table 1. Sample Characteristics by Group and Training Type

		Healthy Controls		AUDs	
		VS training n=40	WM training n=49	VS training n=43	WM training n=39
Age	*M* (*SD*)	21.9 (2.9) [a]	22.6 (2.7) [a]	22.1 (2.3) [a]	21.3 (1.7) [a]
Female N/%	*N* / *%*	N=23 / 58%	N=29 / 59%	N=24 / 56%	N=17 / 44%
Education (years)	*M* (*SD*)	14.5 (1.5) [a]	14.9 (2.0) [a]	14.6 (0.9) [a]	14.5 (1.2) [a]
Drinking Habits					
Frequency (days per wk)	*M* (*SD*)	1.81 (1.2) [b]	1.56 (0.9) [b]	3.95 (1.7) [a]	3.90 (1.6) [a]
Quantity (per occasion)	*M* (*SD*)	2.91 (2.3) [b]	2.72 (2.4) [b]	8.25 (4.5) [a]	9.96 (4.9) [a]
LT SUD Problems					
Alcohol	*M* (*SD*)	2.33 (2.5) [b]	3.4 (5.1) [b]	44.3 (17.8) [a]	40.3 (17.0) [a]
Cannabis	*M* (*SD*)	0.00 [b]	0.35 (1.2) [b]	10.2 (11.2) [a]	10.36 (10.5) [a]
Other drugs	*M* (*SD*)	0.00 [b]	0.00 [b]	8.56 (24.7) [a]	8.77 (15.2) [a]
Training Task Performance		Search Hands Search Letters	OS Task Symmetry Sp	Search Hands Search Letters	OS Task Symmetry Sp
Session 15 average score	*M*(*SD*)	12.42 (2.4) [a]	10.10 (3.2) [b]	12.43 (1.9) [a]	9.40 (3.3) [b]
Highest level achieved	*M*(*SD*)	7.23 (1.1) [a]	6.84 (3.1) [a]	7.03 (0.9) [a]	7.01 (3.2) [a]

AUD = Alcohol Use Disorder; VS = Visual Search adaptive training (control training); WM = Working Memory adaptive training (active training). Drinking Habits = average drinking habits over past 3 months; Frequency = number of drinking days per week; Quantity = average drinks per drinking day. LT SUD Problems = Lifetime Substance Use Disorder problem counts on the SSAGA diagnostic interview; Other drugs refer to drugs other than cannabis and alcohol. Search Hands and Search Letters are the VS training tasks, the listed measure of performance is the average correct set size in the final training session #15. The OS Task (Operation Span task) and Symmetry Sp = (Symmetry Span task) were the WM training tasks, the listed measure of performance is the average score in the final testing session # 15.

There were no significant group or training condition differences in training task performance measures, *p*s between 0.353 – 0.971. Means with a different superscript are significantly different, a > b ($p < .05$).

Figure Captions.

Figure 1. Delay discounting rate (log10 transformed) at baseline (base), 4 weeks (immediate post training), and 8 weeks (4 week post training follow-up) after WM training (WMT) and Visual Search Training (VST).

Figure 2. Interference measures on the Flanker task (Panel A), the Spatial Stroop (Panel B), and Color Stroop (Panel C) at baseline (base), 4 weeks (immediate post training), and 8 weeks (4 week post training follow-up) after WM training (WMT) and Visual Search Training (VST). RT = Reaction Time. Interference was calculated as the RT on incongruent trials minus the RT for neutral trials at each time point for each measure.

Figure 3. Panel A. The average quantity of alcohol consumed per drinking occasion after WM training (WMT) and Visual Search (VS) training at baseline (base), 4 weeks (immediate post training), and 8 weeks (4 week post training follow-up) in Control (con) and AUD (Alcohol Use Disorder Groupts. Panel B changes in the average frequency of drinking alcohol (occasions per week).

Figure 1.

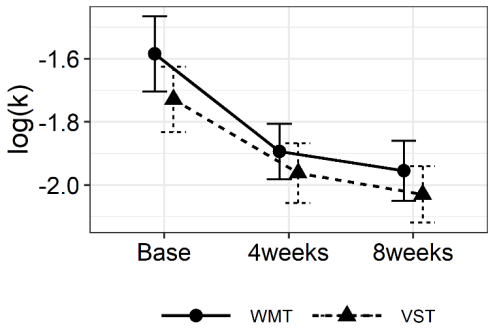


Figure 2.

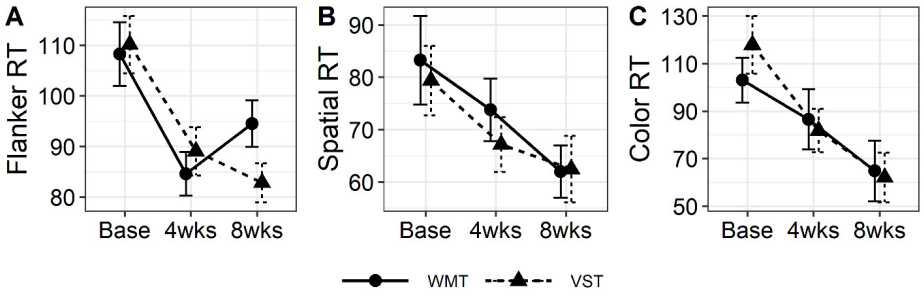
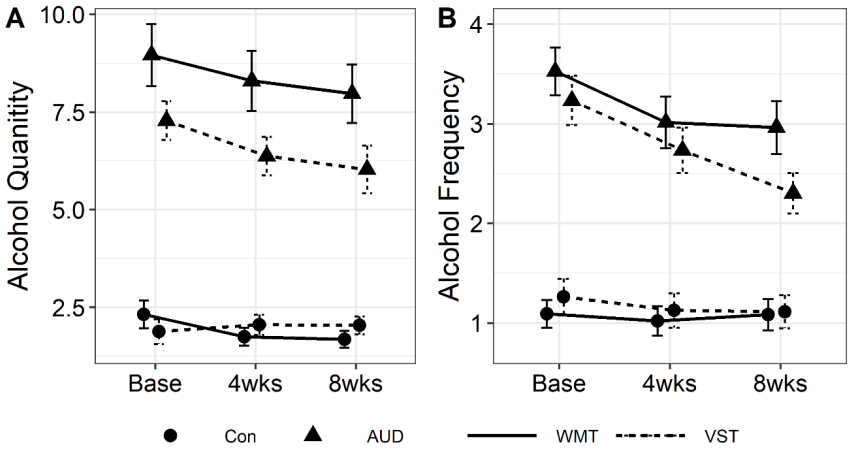

A
Flanker RT
110
100
90
80
Base 4wks 8wks
B
Spatial RT
90
80
70
60
Base 4wks 8wks
C
Color RT
130
110
90
70
50
Base 4wks 8wks
WMT
VST

Figure 3.


A
B
Alcohol Quanitity
Alcohol Frequency
10.0
7.5
5.0
2.5
4
3
2
1
Base
4wks
8wks
Con
AUD
WMT
VST