Predictors of Effective Working Memory Training in Individuals with Alcohol Use Disorders

Rachel L. Gunn, Kyle R. Gerst, Elizabeth A. Wiemers, Thomas S. Redick, and Peter R. Finn

Background: Low working memory (WM) capacity is associated with alcohol use disorders (AUDs). The importance of WM to adaptive functioning has led to a recent influx of studies attempting to improve individual WM capacity using various cognitive training methods. The present study aimed to examine the efficacy of complex WM training for improving WM capacity among individuals with AUD.

Methods: Individuals were randomized to complete either adaptive WM training or active control training. We applied a methodologically rigorous and structured approach, including a battery of near and moderate transfer measures in those with AUDs and a control group. Additionally, we examined cognitive factors (at baseline) and other predictors of adherence, training task improvement, and transfer.

Results: Results suggest improved WM in individuals with AUDs and controls, as evidenced by improved scores on several transfer measures, after adaptive WM training. However, individuals with AUDs showed poorer adherence and less improvement on the training tasks themselves. Neither IQ, WM, sex, nor condition predicted adherence. Level of training task performance, baseline WM, and IQ predicted transfer task improvement.

Conclusions: This is the first study to rigorously examine both the efficacy of WM training in those with AUDs, and predictors of successful training program adherence and transfer in a large sample. Among study completers, results suggest that AUD status does not predict training improvement and transfer. However, AUD status did predict lower program adherence. WM training was more effective in those with higher cognitive ability at baseline. This study provides direct translation to the development of cognitive interventions for treating AUD.

Key Words: Alcohol Use Disorders, Working Memory Training, Executive Cognitive Functioning, Cognitive Training.

POOR SELF-REGULATION AND reduced behavioral control are hallmarks of alcohol use disorder (AUD) and are characterized by impulsive decision making, difficulty inhibiting prepotent responses, and reward biases (Finn, 2002; Iacono et al., 1999). Several core executive cognitive functions have been linked to poor behavioral control and self-regulation (Finn et al., 2015; Hofmann et al., 2009, 2012; de Wit, 2009). In particular, working memory (WM) capacity is a mechanism integral to self-control and decision making (Broadway et al., 2010; Hofmann et al., 2009). Further, research suggests that WM capacity may partially explain the association between externalizing psychopathology, including alcohol problems, and impulsive decision making and behavioral control (Finn et al., 2015; Gunn and Finn, 2013, 2015; Wiers et al., 2010). These findings suggest that those with AUDs serve to benefit from improved WM and self-regulation.

The use of cognitive training methods to improve a variety of individual cognitive functions has received increased attention over the past decade of research. The methods and approaches used in WM training vary widely. Promising research is derived from "core" training or programs thought to target domain-general mechanisms of WM. Characteristically, these programs include multiple tasks spanning modalities, use interference to test maintenance, enforce the use of speed for encoding and retrieval, adapt to individual abilities, minimize automatization, and demand high cognitive workloads or engagement (Morrison and Chein, 2011). Therefore, these programs seek to "train" *the central executive* (Baddeley, 2007), hypothesized to be the core component of WM, responsible for decision making and inhibitory control (Baddeley and Hitch, 1974; Cowan, 1995). Specifically, recent *complex span training* programs require simultaneous storage and processing while performing distracting tasks, thus targeting domain-general attentional control mechanisms of WM (Foster et al., 2017; Harrison et al., 2013). However, a

From the Center for Alcohol and Addiction Studies (RLG), Brown University School of Public Health, Providence, Rhode Island; Department of Psychological and Brain Sciences (KRG, PRF), Indiana University-Bloomington, Bloomington, Indiana; and Department of Psychological Sciences (EAW, TSR), Purdue University, West Lafayette, Indiana.

Received for publication May 2, 2018; accepted September 16, 2018.

Reprint requests: Rachel L. Gunn, PhD, Center for Alcohol and Addiction Studies, Brown University, Box G-S121-4, Providence, RI 02912; Tel.: 401-863-6626; Fax: 401-863-6647; E-mail: rachel_gunn@brown.edu

DOI: 10.1111/acer.13892

significant portion of the WM training literature utilizes *n*-back training programs (e.g., Jaeggi et al., 2008, 2011), which have been repeatedly shown to measure a different construct from complex span tasks (Jaeggi et al., 2010; Kane et al., 2007; Redick and Lindsey, 2013), making it difficult to draw conclusions about transfer to complex WM. Given that complex span measures of WM are related to self-regulation and decision making, these methods provide an appropriate framework for examining WM training in those with AUDs.

Six studies have examined the effects of WM training in alcohol- or drug-use populations: stimulant addicts (Bickel et al., 2011), methamphetamine addicts (Brooks et al., 2017), methadone-maintenance patients (Rass et al., 2015), heavy alcohol users (Houben et al., 2011), alcohol-dependent individuals (Snider et al., 2018), and a sample of inpatients diagnosed with alcohol, cannabis, or cocaine use disorder (Wanmaker et al., 2017). Although several studies provide evidence of far transfer by reduced impulsive decision making on the delay discounting task (Bickel et al., 2011), reduced drinking (Houben et al., 2011), and improved self-regulation (Brooks et al., 2017), they also included a number of methodological limitations. Examples include unstandardized training programs (e.g., variable number of training sessions (Bickel et al., 2011), confounding control conditions (Houben et al., 2011), passive controls (Brooks et al., 2017), assessment of transfer with tasks similar to those trained on (Houben et al., 2011), and use of samples simultaneously enrolled in confounding active treatments (Brooks et al., 2017; Wanmaker et al., 2017). It is essential to first examine whether WM training results in expected improvements in cognitive outcomes (i.e., WM capacity itself), independent of expectancy effects, before it can be tested for clinical outcomes and intervention.

As several reviews have noted (Melby-Lervåg et al., 2016; Shipstead et al., 2012), there is a significant debate regarding the efficacy of WM training, and studies are often conducted with significant methodological concerns, including no or nonactive control groups, narrow outcome measures, small sample sizes, and lack of delayed follow-up sessions. In order to appropriately test the efficacy of a WM training program, an active control group is essential. To create an appropriate control condition, participants should be randomly assigned to either the active WM training, or a control condition that is matched on duration of the intervention, involves similar training procedures (e.g., attending the same number of laboratory sessions), and is adaptive to control for motivation. Extensive batteries of outcome measures that utilize a variety of WM tasks across verbal and visuospatial domains to assess transfer effects further increase robustness of findings. In addition, measuring transfer at a delayed follow-up is essential to show that training effects are maintained after training is complete. Finally, recent work suggests the importance of examining baseline predictors of transfer. For instance, one study suggests that lower baseline cognitive ability (e.g., as may be expected in clinical populations such as those with an AUD) may leave more "room to improve" on WM capacity (Au et al., 2015). However, more recent studies using the same methods employed here suggest that individuals with higher baseline WM may show greater improvement (Foster et al., 2017). This examination of the demographic and cognitive factors associated with WM training-related gains is particularly important for ongoing implementation and dissemination of these methods.

With this in mind, the current study applies a number of important methodological considerations (outlined below in "Materials and Methods") to examine the following: (i) the efficacy of a complex WM training program in those with AUDs; and (ii) several predictors of training improvement, transfer, and adherence. We examined these aims by randomizing individuals with AUDs and controls (individuals without an AUD or any other externalizing disorders), recruited without expectation of improved cognitive capacity, to a complex span WM training program or an active control training condition.

MATERIALS AND METHODS

The present study applied a number of important rigorous methodological considerations, including: (i) an adaptive (complex-span) WM training program (Harrison et al., 2013) that required significant time (15 sessions in less than 4 weeks), rigor, and both verbal and spatial WM tasks; (ii) use of an active and adaptive control training condition (visual search [VS]); (iii) inclusion of a control group; (iv) a battery of near and moderate transfer measures; (v) immediate and delayed (1-month) follow-up assessment; (vi) laboratory-controlled testing conditions; (vii) examination of individual predictors of successful training; and (viii) recruitment of participants without advertisement of potential cognitive benefits (in an effort to reduce expectancy biases; see Foroughi et al., 2016). It should also be noted that the materials presented in the present analyses were a subset of a larger study (in which additional measures were collected) across the testing sessions. As an important first step, we present measures used to focus only on cognitive (i.e., WM) transfer effects. The study was approved by the university's institutional review board.

Participants and Recruitment

Participants (N = 145) met criteria for 1 of 2 groups: those with a current AUD (n = 69) and control participants (n = 76). AUD group status was confirmed using DSM 5 (American Psychiatric Association, 2013) criteria for current AUD. Given the high rates of AUD in the target sample (young adults), individuals were required to meet criteria for a current moderate or severe AUD (presence of 4 or more symptoms) in order to ensure a sample with significant clinical symptomatology. Control participants did not meet criteria for any AUD or other externalizing psychopathology (conduct disorder, antisocial personality disorder, attention-deficit/hyperactivity disorder, cannabis use disorder, or any other substance use disorder). In order to recruit a representative sample, individuals in the AUD group were able to qualify for these other externalizing disorders. Participants were recruited from the community and campus in Bloomington, IN, using advertisements and flyers calling for individuals similar to the target samples (e.g., "impulsive individuals," "heavy drinkers," and "individuals interested in psychological research," "quiet, introspective, and reflective individuals"). Respondents were screened via telephone before being scheduled for the study. Participants in both groups met the following inclusion criteria: (i) between the ages of 18 and 30 years old; (ii) read/speak English; (iii) have at least a sixth-grade education; (iv) have no history of psychosis or head trauma; and (v) 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.

Procedures

Participants completed a total of 19 testing sessions. Two baseline sessions were used to determine full inclusion and group (AUD or control) eligibility and assess baseline cognitive capacity. Following the completion of baseline sessions, participants were scheduled for training sessions. Each participant completed a total of 15 training sessions. Completion of training sessions was immediately followed by the first follow-up session (within 4 days of the final training session). In this session, participants completed the same battery of WM transfer measures administered at baseline. After 1 month (or 30 days) had elapsed from the date training was completed, participants completed a second follow-up session, in which the same battery of WM measures was administered. All sessions were completed in the laboratory, and participants were financially compensated for their participation with $12 hourly for the baseline and follow-up sessions and $15 for each training session. They also had to opportunity to receive the following bonuses: $12 for being ontime to baseline and training sessions, training performance bonuses ranging from $0 to $46, and a $20 completion bonus. Total participant payments ranged from $520 to $870.

Materials

Diagnostic Assessment. The Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA-IV) interview (Collaborative Study on the Genetics of Alcoholism [COGA], 2005) was used to assess group criteria and symptom-level predictors of training efficacy. The following modules were administered: AUD, drug and marijuana use disorder, attention-deficit/hyperactivity disorder, antisocial personality disorder, and conduct disorder.

Active WM Training and VS Control Training. Active training (AT) consisted of the adaptive operation span (OS) and symmetry span (SS) tasks (Harrison et al., 2013). Participants were told they were performing these tasks to improve performance and that the tasks may get easier over time. However, participants were not told that their cognitive capacity was expected to improve in order to offset potential expectancy effects.

In the OS task, participants were presented with a series of trials in which they were asked to make a judgment about accuracy 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 participant improved on the task. All training began at Level 1 difficulty (set sizes 2 to 4), and set size increased by 1 if the participant 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 participant's accuracy on the equations or letters was 75% or less, the level of the next set decreased by 1. In each training session, 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 were asked to make a judgment of symmetry from the vertical axis. Then, participants were presented with a 4×4 matrix with 1 square highlighted in red, which they were instructed to remember. The number of judgments made and matrix positions per level was the same as the OS task, and the task had identical criteria for level progression.

In the control training, adaptive VS tasks (hands and letters) were used. VS tasks are an appropriate control, as they have been shown to be unrelated to measures of WM (Kane et al., 2006) and have been successfully utilized in a number of WM training studies (De Simoni and von Bastian, 2018; Foster et al., 2017; Guye and von Bastian, 2017; Harrison et al., 2013; Redick et al., 2013; von Bastian et al., 2013). As evidenced by others (Foster et al., 2017; Harrison et al., 2013; Redick et al., 2013), these tasks result in practice-related gains, without significant transfer to other measures of individual WM capacity. In these tasks, participants were presented with an array of letters or hands that included a target stimulus, "F" in the letters version and right pointing hand in the hands version, facing either the right or left (mirror-reversed), and an array of distractor letters or hands (normally and mirror-oriented Es/hands) and/or inverted Ts/hands. Each trial began with a fixation dot in the center of the screen, followed by an array of letters/hands for 500 ms, then a mask (16×16 array of black squares) for 2,500 ms. The size of the letter/hand array varied based on difficulty, beginning with a 2×2 array and advancing based on performance to a maximum of 16×16. 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 participant was accurate on 87.5% of the trials and decreased in difficulty if the participant was less than 75% accurate.

WM Transfer Measures. Individual WM capacity was measured with a battery of tasks. These measures were used to assess both baseline WM capacity and training-related gains. Two categories of training effects were measured: near transfer and moderate transfer. The near transfer battery was made of tasks structurally similar to those trained on, but with different stimuli (complex span tasks). The moderate transfer battery was made of WM tasks that are structurally different from those trained on (updating tasks). Performance on each of the following tasks was measured by the number of correct stimuli recalled; higher values reflected higher individual WM capacity.

Near transfer tests included the rotation span (RTS; Harrison et al., 2013), the reading span (RDS; Harrison et al., 2013), and the auditory consonant trigram (ACT; Brown, 1958). In the RTS task, a letter is presented to the participant in either the correct direction or the mirror-reversed direction. The letter is then rotated to 1 of 8 different angles, and participants are asked to determine whether the letter is facing the correct direction or the mirror-reversed direction when turned upright. After making the judgment, participants see a short or long arrow pointing in 1 of the 8 directions. Then, after a set of judgments followed by arrows, participants are asked to click on the arrows that were presented in the correct order. There were a total of 15 trials, with 3 to 10 arrows each. In the RDS, participants are presented with a series of sets of trials (ranging from 3 to 7 trials), which involve reading a grammatically correct sentence that either does or does not make logical sense. Participants respond "TRUE" or "FALSE" and are immediately presented with a letter to keep in mind. At the end of each set, they recall the letters in the correct order. In the ACT, participants are presented with 3, 4, and 5 nonsense consonant strings, and then asked to count backward by 3s from a 3-digit number for 18 or 36 seconds. As in the other near transfer tasks, they are asked to recall the consonant strings after each set of trials. Each of these tasks (RTS, RDS, and ACT) taps the attentional control mechanisms of the central executive of WM capacity by requiring the participant to retain information while completing secondary distracting tasks.

The *moderate transfer* effects of training on WM capacity were assessed using the running letter span (RLS; Harrison et al., 2013), the running spatial span (RSS; Harrison et al., 2013), and the keep track (KT; Harrison et al., 2013) tasks. In the RLS task, a series of letters were presented 1 at a time (2 letters per second). Before the presentation, participants were told they are to remember a set size (number of letters to recall) and recall the most recent set size (n) in the correct serial position. For example, participants could have been told to remember the last 5 letters shown and then presented with 7 letters. Set sizes vary from 3 to 9, and 2 trials per set size were completed. The RSS task is identical to the RLS task, except that to-be-remembered stimuli are matrix locations on a 4 × 4 matrix. In the KT task, participants are presented sequentially with 16 words from 6 categories. They are told to remember the most recent instances of a certain number of categories. There were 15 trials, and the number of to-be-remembered categories per trial (set size) was 2, 3, 4, 5, or 6. Before each trial, participants are told the category of the most recent instance they should recall. Each of these tasks requires the participant to maintain a string of stimuli and respond based on a varying set size. It should be noted that in Harrison and colleagues (2013) these tasks were also referred to as near transfer measures. However, here we have reclassified them as moderate transfer measures given they are structurally different from the complex span tasks (e.g., no secondary processing task). Although both types of tasks measure WM, improvement on theses updating tasks would represent stronger evidence that individual core WM capacity has been altered, given their unique measurement characteristics and the lower likelihood that strategies practiced on the complex span training would as easily transfer to these tasks.

Intelligence. The Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999) was used to assess verbal and performance IQ at baseline. This assessment includes 4 subtests: block design, vocabulary, matrix reasoning, and similarities.

Data Analyses

R (R Core Team, 2013) was used to conduct all analyses. In order to examine training task improvement and transfer effects, *lme4* (Bates et al., 2015) was used to perform linear mixed-effects analysis. Analysis of variance (ANOVA) and chi-square tests were used to examine differences in demographic characteristics and baseline WM capacity by group and condition. Chi-square and logistic regression analyses were used to examine demographic (sex, group, condition) predictors of training adherence in the full sample of individuals who were assigned to training (n = 189). In order to account for shared variance across time points (baseline, follow-up 1, and follow-up 2), linear mixed models were used to examine the main effects of near transfer and moderate transfer (change in individual WM scores from baseline to follow-up sessions), as well as predictors of these transfer effects in the final sample of individuals who completed training and follow-up sessions (n = 145). Model specifications are described in more detail below for each model.

RESULTS

Final Sample Demographics and Training Program Adherence

A total of 189 individuals were eligible for the study (i.e., met full inclusion criteria after being interviewed using the SSAGA; COGA, 2005). Of those 189, 97 were randomized into the AT and 92 into the VS condition. Seventy-five (77%) participants assigned to the AT completed the full protocol (AUD = 35, control = 40). Seventy (76%) participants assigned to the VS (control) training completed the full protocol (AUD = 34, control = 36), resulting in a full sample of 145. Table 1 presents the final sample demographics (N = 145), as well as split by group and condition. Table 1 also presents ANOVA and chi-square statistics comparing demographic variables by group and condition. There was no significant difference in participants' age, sex, ethnicity, or student status as a function of condition (VS vs. AT), suggesting appropriate baseline comparisons. Additionally, there were no significant differences between group (AUD vs. control) on age, ethnicity, or student status. Among individuals in the AUD group, there were a number of comorbid current externalizing diagnoses: 17% with attention-deficit/hyperactivity disorder, 26% with conduct disorder, 17% with antisocial personality disorder, 18% with cannabis use disorder, and 19% with another substance use disorder.

A chi-square test examining program adherence (successful completion of all 19 sessions including baseline, training, and follow-up sessions) revealed neither gender (χ^2 (1) = 0.06, p = 0.81) nor condition (χ^2 (1) = 0.001, p = 0.98) predicted training completion. Additionally, linear regression revealed neither IQ (B = −0.004, p = 0.07) nor baseline WM (B = 0.0069, p = 0.76) predicted program adherence. However, logistic regression revealed that group did predict adherence (B = 0.25, p < 0.001), in that AUDs had a higher likelihood of dropping from the study before completion.

Efficacy of WM Training: Training Task Improvement and Transfer

Training Task Improvement. A random-intercepts mixed-effects model was applied, using normalized (z-score) session averages, to examine level of improvement on each

Table 1. Demographic Characteristics of Final Sample by Group and Condition

		Group			Condition		
	Full (n = 145)	Controls (n = 76)	AUDs (n = 69)		VS (n = 70)	Active (n = 75)	
Age (M, SD)	22.06 (2.42)	22.30 (2.64)	21.80 (2.14)	F(1, 143) = 1.59, p = 0.21	22.04 (2.63)	22.08 (2.22)	F(1, 143) = 0.01, p = 0.93
% Female	60	67	52	X^2 (1) = 3.24, p = 0.20	66	55	X^2 (1) = 1.73, p = 0.42
% Caucasian	74	70	80	X^2 (1) = 1.40, p = 0.24	76	73	X^2 (1) = 0.02, p = 0.89
% Student	83	84	81	X^2 (1) = 0.07, p = 0.79	86	80	X^2 (1) = 0.48, p = 0.49
IQ (M, SD)	115.42 (12.61)	116.04 (8.58)	114.74 (15.92)	F(1, 143) = 0.38, p = 0.54	114.67 (8.75)	116.11 (15.36)	F(1, 143) = 0.47, p = 0.50

*0.05, **0.01, ***0.001.

Table 2. Random-Intercepts Mixed-Effects Model of Training Task Improvement

Predictors	B (95% CI)
(Intercept)	**−1.23 (−1.43, −1.04), $p < 0.001$**
Session	**0.17 (0.16, 0.17), $p < 0.001$**
Group	−0.10 (−0.36, 0.17), $p = 0.488$
Condition	**0.53 (0.26, 0.79), $p < 0.001$**
Session × group	**−0.01 (−0.02, −0.00), $p < 0.01$**
Group × condition	−0.03 (−0.33, 0.39), $p = 0.867$
Session × condition	**−0.07 (−0.08, −0.06), $p < 0.001$**

$\sigma^2 = 0.328$, Subject ICC = 0.522; group = AUD vs. controls; condition = active WM training vs. visual search (control) training; session = Training sessions 1 to 15. Bold values are significant at $p < 0.05$ or lower, as presented for each parameter.

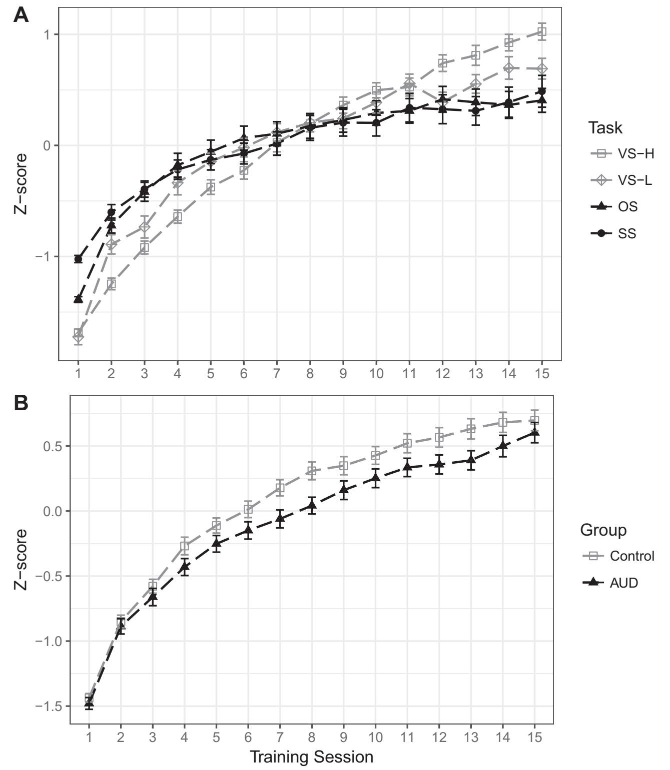

Fig. 1. Average training task (*z*-score) for each training task (**A**) and averaged across group (**B**) at each training session. Control training tasks are represented by VS-H = visual search-hands, VS-L = visual search-letters, and active training tasks are represented by OS = operation span, SS = symmetry span.

task. In order to examine effects of condition and group in the same model, normalized *z*-scores were computed for a combined score of training tasks received (AT or VS) for each participant. Fixed effects of session (1 to 15), condition (AT and VS), group (AUD and control), and interactions of session by group, session by condition, and group by condition were entered into the model, along with random intercepts of participant and task. Full results from this analysis are reported in Table 2. This model revealed main effects of session, indicating improvement on training tasks as session increased, and session-by-condition interaction, suggesting those in VS improved more than those in the AT (Fig. 1*A*). This model also revealed a session-by-group interaction, suggesting controls improved more across training sessions compared to AUDs (Fig. 1*B*). However, a group-by-condition interaction was not significant, suggesting that the groups performed similarly across conditions.

Transfer Effects. ANOVA was used to examine baseline differences in transfer measures between group and condition. Importantly, there were no significant differences between transfer measures by randomized conditions for 5 of 6 measures at baseline, suggesting appropriate comparisons for those 5 tasks at follow-up (RDS: $F(1, 140) = 1.69$, $p = 0.19$; RTS: $F(1, 139) = 0.66$, $p = 0.42$; ACT: $F(1, 142) = 0.01$, $p = 0.91$; RLS: $F(1, 140) = 0.12$, $p = 0.73$; RSS: $F(1, 140) = 1.21$, $p = 0.27$; KT: $F(1, 140) = 5.02$, $p = 0.03$). Additionally, there were no baseline differences on any WM transfer measures between groups (RDS: $F(1, 140) = 0.96$, $p = 0.33$; RTS: $F(1, 142) = 0.06$, $p = 0.81$; ACT: $F(1, 142) = 0.001$, $p = 0.97$; RLS: $F(1, 140) = 1.94$, $p = 0.16$; RSS: $F(1, 140) = 0.05$, $p = 0.83$; KT: $F(1, 140) = 0.55$, $p = 0.46$). Linear mixed models were run for each transfer measure separately to examine effect of session (baseline, follow-up 1, and follow-up 2), condition (AT and VS), group (AUD and control), and group-by-condition, session-by-condition, and session-by-group interactions, with random intercepts of subject and session. Full models are presented in Table 3.

Near Transfer—Models testing transfer effects revealed significant session-by-condition interactions on 2 near transfer measures, RTS and ACT, at initial follow-up, reflected by improved scores for those subjects randomized into AT. At follow-up 2 (1 month), both effects remained significant, suggesting maintenance of transfer effects 1 month after training completion. These results are presented in Fig. 2*B*,*C*. ACT also revealed a session-by-group interaction at follow-up 2, indicating controls had higher scores at follow-up 2, compared to the AUD group. However, this effect is nonsignificant at follow-up 1. The third near transfer task (RDS) revealed no evidence of transfer.

Moderate Transfer—Additionally, a significant session-by-condition interaction was observed on the RSS task, suggesting transfer by improved scores for those subjects randomized into AT (Fig. 2*E*). This effect was not maintained at follow-up session 2. There was no evidence of transfer for the other 2 moderate transfer measures (RLS or KT). Table 3 and Fig. 2*D*–*F* present all moderate transfer results, including null findings.

Predictors of Training Improvement and Transfer Task Performance

Predictors of Training Task Improvement. In order to more closely examine improvement on the training tasks themselves, cognitive predictors of training improvement

Table 3. Linear Mixed Model of Near and Moderate Transfer Effects at Follow-Up 1 and Follow-Up 2

	B (95% CI)					
	RDS	RTS	ACT	RLS	RSS	KT
Group	−4.56 (−12.23, 3.11)	−2.33 (−8.31, 3.65)	−1.28 (−5.06, 2.49)	**−5.82 (−11.50, −0.15)***	−2.43 (−7.58, 2.73)	−1.94 (−5.40, 1.52)
Condition	−6.05 (−13.67, 1.57)	−4.14 (−10.06, 1.77)	−1.13 (−4.88, 2.62)	−4.52 (−10.15, 1.12)	−3.93 (−9.04, 1.18)	**−3.67 (−7.11, −0.23)***
Group × condition	5.59 (−4.20, 15.37)	5.05 (−2.63, 12.72)	2.76 (−2.26, 7.78)	7.12 (−0.40, 14.63)	4.28 (−2.32, 10.87)	2.58 (−1.85, 7.00)
F1 × condition	−4.19 (−10.31, 1.92)	**−7.79 (−12.37, −3.20)*****	**−2.85 (−5.33, −0.37)***	−2.69 (−6.49, 1.12)	**−4.47 (−8.54, −0.40)***	1.51 (−1.22, 4.24)
F2 × condition	−4.23 (−10.30, 1.84)	**−5.69 (−10.26, −1.11)***	**−4.05 (−6.53, −1.58)****	1.62 (−2.18, 5.43)	−2.25 (−6.30, 1.81)	1.08 (−1.65, 3.80)
F1 × group	−0.16 (−6.28, 5.96)	1.03 (−5.63, 3.56)	−1.88 (−4.36, 0.60)	−0.35 (−4.16, 3.46)	−3.06 (−7.14, 1.01)	−2.30 (−5.03, 0.43)
F2 × group	−0.54 (−6.62, 5.53)	−3.33 (−7.91, 1.62)	**−3.73 (−6.20, −1.26)****	−2.93 (−6.73, 0.88)	**−4.74 (−8.81, −0.68)***	**−5.16 (−7.89, −2.44)*****

Effects of interest displayed, full model description provided in text.
ACT, auditory consonant trigram; F1, follow-up 1; F2, follow-up 2; KT, keep track task; RDS, reading span; RTS, rotation span; RLS, running letter span; RSS, running spatial span; group = AUD vs. controls, condition = active WM training vs. visual search (control) training; session = baseline, follow-up 1, follow-up 2.
* 0.05, **0.01, *** 0.001.

were considered. A change score was used for each training task to examine whether cognitive (baseline WM and IQ) factors predicted improvement on any of the 4 training tasks (OS, SS, visual search-hands [VS-H], or visual search-letters [VS-L]). Each change score was calculated as the difference between average level reached on session 15 and average level reached on session 1. As presented in Table 4, linear regression models (controlling for group) revealed that baseline WM (measured by the OS task) did not predict improvement on any of the 4 tasks. However, baseline IQ (measured by the WASI) did predict change in both AT tasks (OS and SS), but neither of the VS (control) training tasks. In other words, higher IQ was associated with more training gains on the AT tasks, but not the VS tasks. Group effects were nonsignificant in all models.

Training Tasks Performance as Predictor of Transfer. In order to consider level of improvement on trained tasks as a predictor of transfer task improvement, average score at the final session (training session 15) was entered into a random-intercepts linear mixed models as a fixed effect for each training task. In each model, session, group, and IQ were also considered fixed effects, and session and subject variables were random. Results are presented in Table 5. These analyses revealed a significant effect of final training score on transfer on all near and moderate transfer measures in both the OS and SS training tasks. However, analyses of performance on the VS tasks did not predict transfer task performance on most tasks. The final session performance in the VS-H task only predicted transfer task improvement on 1 task (RDS). IQ was nonsignificant in all models except the ACT for both the VS-L ($B = 0.28$, $p = 0.01$) and VS-H ($B = 0.30$, $p = 0.003$) models. Similar to main effect transfer models, group effects remained nonsignificant in all near transfer models. Regarding moderate transfer models, the effect of group was significant on the RLS task in the OS model ($B = -5.84$, $p = 0.02$) only. Overall, these results suggest that actual performance on the AT tasks was an important predictor of transfer task performance, regardless of baseline IQ, and suggests that completion of the AT had a larger impact on transfer measures than the VS (control) training.

Cognitive Predictors of Transfer Task Improvement. Linear mixed models were used again to examine whether cognitive ability (WM and IQ) at baseline predicted greater improvement from baseline to follow-up 1 and follow-up 2 on transfer tasks. Again, session and subject were set as random intercepts. Other fixed effects included session, condition, baseline WM/IQ, and group. Results are presented in Table 6. Baseline WM (OS) predicted greater improvement on all 3 near transfer tasks and 2 moderate transfer tasks (RLS and KT). Additionally, baseline IQ (WASI) predicted greater improvement on 2 near transfer measures (RTS and ACT) and all 3 moderate transfer measures. Group effects were nonsignificant in all near transfer models. However, group effects were significant in both RLS models examining

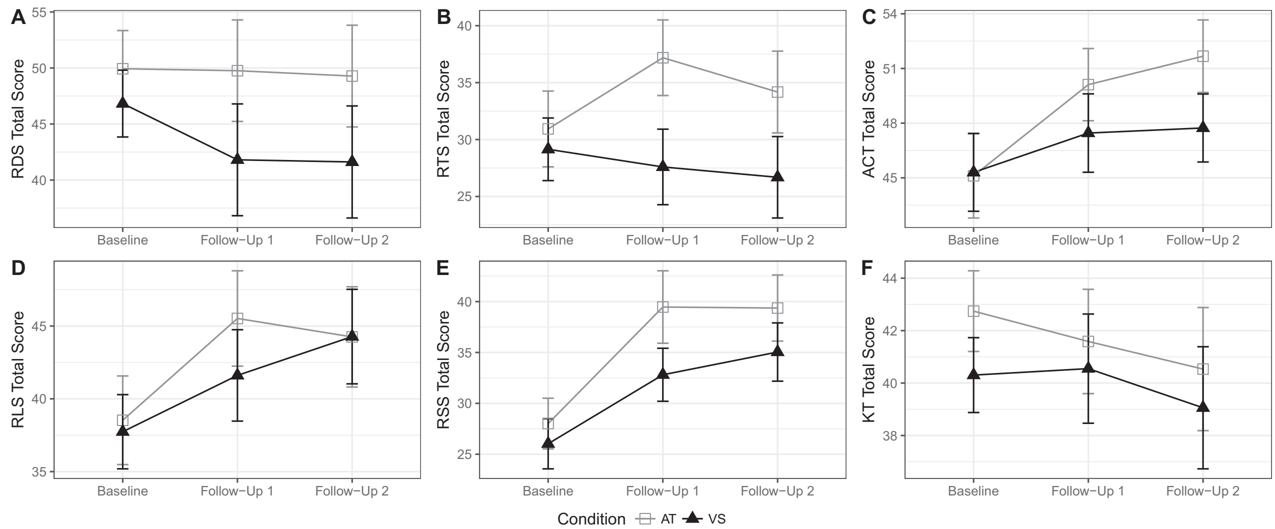

Fig. 2. Figure displays average total score for each near (**A**–**C**) and moderate (**D**–**F**) transfer measures at each session by condition across both groups. ACT, auditory consonant trigram; RDS, reading span; RTS, rotation span; RLS, running letter span; RSS, running spatial span; KT, keep track task; AT, active training condition; VS, visual search condition. Error bars represent 95% CI.

Table 4. Linear Regression, Cognitive Predictors of Training Task Improvement

	O-Span change	S-Span change	VS-H change	VS-L change
Predictor	B (95% CI)			
WM	−0.00 (−0.05, 0.05)	−0.03 (−0.08, 0.02)	−0.00 (−0.04, 0.03)	0.01 (−0.01, 0.02)
IQ	**0.07 (0.03, 0.11)****	**0.09 (0.05, 0.13)*****	−0.02 (−0.03, 0.07)	−0.00 (−0.03, 0.03)

Effects of interest displayed, full model description provided in text.
ACT, auditory consonant trigram; KT, keep track task; RDS, reading span; RTS, rotation span; RLS, running letter span; RSS, running spatial span.
*0.05, **0.01, ***0.001.

Table 5. Linear Mixed Model, Final Active Training Performance Predicting Transfer Effects

	RDS	RTS	ACT	RLS	RSS	KT
Session 15	B (95% CI)					
OS	**2.10 (1.16, 3.04)*****	**1.00 (0.12, 1.81)***	**0.64 (0.12, 1.16)***	**1.71 (0.97, 2.44)*****	**1.48 (0.71, 2.26)*****	**0.95 (0.50, 1.40)*****
SS	**2.27 (1.18, 3.35)*****	**1.75 (0.87, 2.64)*****	**0.66 (0.06, 1.26)***	**1.58 (0.70, 2.47)*****	**2.26 (1.49, 3.04)*****	**0.88 (0.34, 1.41)****
VS-L	1.57 (−2.54, 5.68)	1.84 (−1.38, 5.05)	1.13 (−0.45, 2.70)	0.68 (−2.36, 3.72)	−0.15 (−2.74, 2.43)	0.40 (−1.49, 2.28)
VS-H	**2.39 (0.38, 4.40)***	1.07 (−0.58, 2.71)	0.40 (−0.42, 1.22)	0.31 (−1.24, 1.87)	1.18 (−0.11, 2.46)	−0.05 (−1.02, 0.92)

Effects of interest displayed, full model description provided in text.
ACT, auditory consonant trigram; RDS, reading span; RTS, rotation span; RLS, running letter span; RSS, running spatial span; KT, keep track task; VS-H, visual search-hands; VS-L, visual search-letters; OS, operation span; SS, symmetry span.
Bolded *p*-values <0.05,*0.05, **0.01, ***0.001.

WM ($B = -4.24$, $p = 0.05$) and IQ ($B = -4.45$, $p = 0.04$), and the KT model examining WM ($B = -3.07$, $p = 0.02$), indicating those with AUD were less likely to show improvement when considering cognitive predictors.

DISCUSSION

This study investigated the effects of complex WM training on individuals with and without AUDs. Consistent with previous studies examining complex span training (Foster et al., 2017; Harrison et al., 2013), results revealed significant transfer on 2 near WM transfer measures (RTS and ACT) at immediate follow-up only for individuals who completed the AT. Importantly, these effects were also maintained at the second follow-up session and were not moderated by group in main effect models. Further, and novel to our study, we also found evidence of transfer on 1 moderate transfer task (RSS), also independent of group. However, this transfer effect was not maintained at 1-month follow-up. We also found that baseline cognitive ability appeared to be

Table 6. Linear Mixed Models, Cognitive Predictors of Transfer Tasks

	RDS	RTS	ACT	RLS	RSS	KT
Predictors	B (95% CI)					
WM	**0.34 (0.17, 0.50)*** **	**0.28 (0.14, 0.41)*** **	**0.14 (0.06, 0.23)*** **	**0.25 (0.12, 0.38)*** **	0.08 (−0.04, 0.19)	**0.13 (0.05, 0.20)** **
IQ	0.19 (−0.00, 0.38)	**0.18 (0.03, 0.33)***	**0.18 (0.08, 0.28)*** **	**0.19 (0.04, 0.33)***	**0.21 (0.08, 0.34)** **	**0.17 (0.09, 0.25)*** **

Effects of interest displayed, full model description provided in text.
ACT, auditory consonant trigram; KT, keep track task; OS, operation span; RDS, reading span; RTS, rotation span; RLS, running letter span; RSS, running spatial span; SS, symmetry span; VS-H, visual search-hands; VS-L, visual search-letters.
Bolded *p*-values <0.05, *0.05, **0.01, ***0.001.

important predictors of WM training and transfer task improvement. These results suggest that among individuals who completed training, training gains are at least as likely in individual with AUDs, compared to control subjects. Further, these improvements in WM occurred in a sample of individuals with AUD and rates of comorbid externalizing disorders that represent of the population at large (Compton et al., 2005; Kessler et al., 2006; Stinson et al., 2005), increasing the generalizability of our findings.

In addition to examining the efficacy of WM training in individuals with AUDs, we also sought to identify predictors of program adherence, training task improvement, and transfer. To our knowledge, no studies have investigated the factors that may be associated with cognitive training program adherence among clinical samples, and very few have examined these factors in healthy populations (Jaeggi et al., 2014). Given the large time commitment and cognitive challenges associated with these protocols, this is an important variable to consider in the potential application of these programs. Importantly, the rate of dropout was higher among individuals with AUDs. There are several factors that may have made it more difficult for those with AUDs to complete the study. For instance, the training schedule required subjects to come in at least 4 times per week, to provide a breath alcohol level of 0.00, and not be experiencing any hangover symptoms. For some heavy problematic drinkers, it may have been difficult to maintain these study requirements. Importantly, neither gender, baseline cognitive ability, nor condition predicted training program adherence.

The analyses presented here included all subjects who completed training, regardless of how much they improved on training tasks. However, unknown factors such as poor motivation, strategy, or additional cognitive skills may have impacted how much individuals improved on the training. Analyses examining the effect of training task performance on transfer tasks support this notion. Notably, final level on the AT tasks predicted improvement on transfer tasks on all 6 (near and moderate) measures. This pattern was only observed in 1 transfer measure for 1 control training task. In other words, those who performed better on the AT tasks themselves were much more likely to see general improvements on individual WM capacity. This is a very important finding and may be a more sensitive way to assess actual transfer to core WM resources. These findings are particularly important when communicating the efficacy of WM training, and discussion of these findings should reflect that the more gains in training, the more likely individuals are to see transfer to core WM capacity.

It is also important to note that baseline intelligence predicted level of training improvement on both of the AT tasks, in that higher intelligence was associated with more training success. Similarly, and consistent with Foster and colleagues (2017), higher baseline WM capacity and IQ predicted improved WM capacity on several near and moderate transfer measures, despite having statistically more "room to improve," as suggested by Au and colleagues (2016). As recently discussed by Redick and colleagues (2015), these findings may predict some barriers for further implementation of WM training methods among those with lower WM and related executive functioning, in this case for individuals with AUDs. Additionally, in these models an effect of group was revealed for 2 moderate transfer tasks, suggesting that when considering baseline cognitive ability, individuals with AUD are less likely to show improvement on transfer measures. However, these effects should be understood in the context of some limited generalizability, given it was not consistent across all tasks (including no significant effects on near transfer tasks) and other models (in which baseline cognitive ability was not accounted for).

It is important to understand the findings of this research within the context of its limitations. In particular, characteristics of the sample limit the generalizability of the results. First, our sample is largely comprised of college students, with a mean IQ of 116, which reflects a higher average than the population at large and represents less variability in cognitive skills. Although our sample size was sufficient for the analyses presented here, we were limited in our ability to examine the role of comorbid externalizing problems on WM improvements. Future studies should recruit larger samples in order to examine the influence of comorbidities on training effects. Additionally, the study was conducted in an on-campus laboratory, which makes access to facilities convenient and familiar and may have affected retention in the positive direction. If these methods were to be implemented in community or medical settings, more barriers to adherence may exist. Finally, it is important to note that there was significant monetary incentive in this study. Individuals who completed the study made an average of

$699.55. The study design did not direct recruitment to individuals looking to improve cognitive or memory skills, as in many cognitive training studies. In fact, a recent study indicates that using suggestive recruitment based on the potential benefits of cognitive training leads to placebo effects on measures of fluid intelligence (Foroughi et al., 2016). In order to complete a rigorous scientific design, we found it important to minimize expectancy effects that may exist in a sample who is actively trying to improve their cognitive skills. However, without this internal motivation, monetary incentives were necessary to motivate improvement on the training and maintain retention of subjects. Future studies should recruit individuals who were specifically motivated to improve cognitive skills and compare them to a "blind" sample in order to understand whether motivation plays an important role in training completion and transfer.

In summary, results suggest that those with AUDs who were able to complete a rigorous protocol improve on WM training programs and show evidence of near and some moderate transfer. However, individuals with AUDs are also less likely to adhere to training program requirements. Further, those who showed more training improvement on the active WM training program were also more likely to show improvement on transfer effects. Results also suggest that individuals who begin the training with higher cognitive ability are also more likely to benefit from training. This and future studies provide direct translation to the development of cognitive interventions for treating AUDs and related externalizing psychopathology.

ACKNOWLEDGMENTS

The authors gratefully acknowledge Kali Falnes, Emily Atkinson, Taylor Daniel, Emily Grayson, Pauline Panayi, and Abigail Berry for their contribution to the project.

FUNDING

This research was primarily supported by NIAAA R01 AA13650 (PI: Finn). Manuscript preparation was supported in part by T32 AA07462 (RLG, KRG), T32 AA007459 (RLG), and NSF Award # 1632403 (TSR, EAW).

CONFLICT OF INTEREST

The authors have no conflicts of interest to declare.

REFERENCES

Au J, Buschkuehl M, Duncan GJ, Jaeggi SM (2016) There is no convincing evidence that working memory training is NOT effective: a reply to Melby-Lervåg and Hulme (2015). Psychon Bull Rev 23:331–337.

Au J, Sheehan E, Tsai N, Duncan GJ, Buschkuehl M, Jaeggi SM (2015) Improving fluid intelligence with training on working memory: a meta-analysis. Psychon Bull Rev 22:366–377.

Baddeley AD (2007) Working Memory, Thought, and Action. Oxford Psychology Series. Oxford University Press, New York.

Baddeley AD, Hitch G (1974) Working memory. Psychol Learn Motiv 8:47–89.

Bates D, Mächler M, Bolker BM, Walker SC (2015) Fitting linear mixed-effects models using lme4. J Stat Softw 67:1–48.

Bickel WK, Yi R, Landes RD, Hill PF, Baxter C (2011) Remember the future: working memory training decreases delay discounting among stimulant addicts. Biol Psychiatry 69:260–265.

Broadway JM, Redick TS, Engle RW (2010) Working memory capacity: self-control is (in) the goal, in *Self Control in Society, Mind, Brain* (Hassin YTR, Ochsner KN eds), pp 163–173. Oxford University Press, New York.

Brooks SJ, Wiemerslage L, Burch K, Maiorana S, Cocolas E, Schiöth H, Kamaloodien K, Stein D (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.

Brown J (1958) Some tests of the decay theory of immediate memory. Q J Exp Psychol 10:12–21.

Collaborative Study on the Genetics of Alcoholism (COGA) (2005) Semi-Structured Interview for the Genetics of Alcoholism-Version Four (SSAGA-IV). Available at: https://niaaagenetics.org/coga_instruments/resources.html. Accessed January 2, 2014.

Compton WM, Conway KP, Stinson FS, Colliver JD, Grant BF (2005) Prevalence, correlates, and comorbidity of DSM-IV antisocial personality syndromes and alcohol and specific drug use disorders in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions. J Clin Psychiatry 66:677–685.

Cowan N (1995) Attention and Memory: An Integrated Framework. Oxford University Press, Oxford.

Finn PR (2002) Motivation, working memory, and decision making: a cognitive-motivational theory of personality vulnerability to alcoholism. Behav Cogn Neurosci Rev 1:183–205.

Finn PR, Gunn RL, Gerst KR (2015) The effects of a working memory load on delay discounting in those with externalizing psychopathology. Clin Psychol Sci 3:202–214.

Foroughi CK, Monfort SS, Paczynski M, McKnight PE, Greenwood PM (2016) Placebo effects in cognitive training. Proc Natl Acad Sci U S A 113:7470–7474.

Foster JL, Harrison TL, Hicks KL, Draheim C, Redick TS, Engle RW (2017) Do the effects of working memory training depend on baseline ability level? J Exp Psychol Learn Mem Cogn 43:1677–1689.

Gunn RL, Finn PR (2013) Impulsivity partially mediates the association between reduced working memory capacity and alcohol problems. Alcohol 47:3–8.

Gunn RL, Finn PR (2015) Applying a dual process model of self-regulation: the association between executive working memory capacity, negative urgency, and negative mood induction on pre-potent response inhibition. Personality Individ Differ 75:210–215.

Guye S, von Bastian CC (2017) Working memory training in older adults: Bayesian evidence supporting the absence of transfer. Psychol Aging 32:732–746.

Harrison TL, Shipstead Z, Hicks KL, Hambrick DZ, Redick TS, Engle RW (2013) Working memory training may increase working memory capacity but not fluid intelligence. Psychol Sci 24:2409–2419.

Hofmann W, Friese M, Strack F (2009) Impulse and self-control from a dual-systems perspective. Pers Psychol Sci 4:162–176.

Hofmann W, Schmeichel BJ, Baddeley AD (2012) Executive functions and self-regulation. Trends Cogn Sci 16:174–180.

Houben K, Wiers RW, Jansen A (2011) Getting a grip on drinking behavior. Psychol Sci 22:968–975.

Iacono WG, Carlson SR, Taylor J, Elkins IJ, McGue M (1999) Behavioral disinhibition and the development of substance-use disorders: findings from the Minnesota Twin Family Study. Dev Psychopathol 11:869–900.

Jaeggi SM, Buschkuehl M, Jonides J, Perrig WJWJ, Buschkeuhl M, Jonides J, Perrig WJWJ (2008) Improving fluid intelligence with training on working memory. Proc Natl Acad Sci U S A 105:6829–6833.

Jaeggi SM, Buschkuehl M, Jonides J, Shah P (2011) Short- and long-term benefits of cognitive training. Proc Natl Acad Sci U S A 108:10081–10086.

Jaeggi SM, Buschkuehl M, Perrig WJ, Meier B (2010) The concurrent validity of the *N*-back task as a working memory measure. Memory 18:394–412.

Jaeggi SM, Buschkuehl M, Shah P, Jonides J (2014) The role of individual differences in cognitive training and transfer. Mem Cognit 42:464–480.

Kane MJ, Conway ARA, Miura TK, Colflesh GJH (2007) Working memory, attention control, and the n-back task: a question of construct validity. J Exp Psychol Learn Mem Cogn 33:615–622.

Kane MJ, Poole BJ, Tuholski SW, Engle RW (2006) Working memory capacity and the top-down control of visual search: exploring the boundaries of executive attention. J Exp Psychol Learn Mem Cogn 32:749–777.

Kessler RC, Adler L, Barkley R, Biederman J, Conners CK, Demler O, Faraone SV, Greenhill LL, Howes MJ, Secnik K, Spencer T, Ustun TB, Walters EE, Zaslavsky AM (2006) The prevalence and correlates of adult ADHD in the United States: results from the National Comorbidity Survey Replication. Am J Psychiatry 163:716–723.

Melby-Lervåg M, Redick TS, Hulme C (2016) Working memory training does not improve performance on measures of intelligence or other measures of "far transfer." Pers Psychol Sci 11:512–534.

Morrison AB, Chein JM (2011) Does working memory training work? The promise and challenges of enhancing cognition by training working memory. Psychon Bull Rev 18:46–60.

R Core Team (2013) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rass O, Schacht RL, Buckheit K, Johnson MW, Strain EC, Mintzer MZ (2015) A randomized controlled trial of the effects of working memory training in methadone maintenance patients. Drug Alcohol Depend 156:38–46.

Redick TS, Lindsey DRB (2013) Complex span and n-back measures of working memory: a meta-analysis. Psychon Bull Rev 20:1102–1113.

Redick TS, Shipstead Z, Harrison TL, Hicks KL, Fried DE, Hambrick DZ, Kane MJ, Engle RW (2013) No evidence of intelligence improvement after working memory training: a randomized, placebo-controlled study. J Exp Psychol Gen 142:359–379.

Redick TS, Shipstead Z, Wiemers EA, Melby-Lervåg M, Hulme C (2015) What's working in working memory training? An educational perspective Educ Psychol Rev 27:617–633.

Shipstead Z, Redick TS, Engle RW (2012) Is working memory training effective? Psychol Bull 138:628–654.

de Simoni C, von Bastian CC (2018) Working memory updating and binding training: Bayesian evidence supporting the absence of transfer. J Exp Psychol Gen 147:829–858.

Snider SE, Deshpande HU, Lisinski JM, Koffarnus MN, Laconte SM, Bickel WK (2018) Working memory training improves alcohol users' episodic future thinking: a rate-dependent analysis. Biol Psychiatry CNNI 3:160–167.

Stinson FS, Grant BF, Dawson DA, Ruan WJ, Huang B, Saha T (2005) Comorbidity between DSM-IV alcohol and specific drug use disorders in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions. Drug Alcohol Depend 80:105–116.

von Bastian CC, Langer N, Jäncke L, Oberauer K (2013) Effects of working memory training in young and old adults. Mem Cognit 41:611–624.

Wanmaker S, Leijdesdorff SMJ, Geraerts E, van de Wetering BJM, Renkema PJ, Franken IHA (2017) The efficacy of a working memory training in substance use patients: a randomized double-blind placebo-controlled clinical trial. J Clin Exp Neuropsychol 40:473–486.

Wechsler D (1999) Manual for the Wechsler Abbreviated Scale of Intelligence. The Psychological Corporation, San Antonio, TX.

Wiers RW, Ames SL, Hofmann W, Krank M, Stacy AW (2010) Impulsivity, impulsive and reflective processes and the development of alcohol use and misuse in adolescents and young adults. Front Psychol 1:144.

de Wit H (2009) Impulsivity as a determinant and consequence of drug use: a review of underlying processes. Addict Biol 14:22–31.