Integrative data analysis of self-efficacy in four clinical trials for alcohol use disorder

Eric S. Kruger, PhD[1], **Kelsey N. Serier, PhD**[1], **Rory A. Pfund, PhD**[1], **James R. McKay, PhD**[2,3], **Katie Witkiewitz, PhD**[1]

[1]University of New Mexico, Albuquerque NM, USA

[2]University of Pennsylvania, Philadelphia, PA, USA

[3]Philadelphia VA Medical Center, Philadelphia, PA, USA

Abstract

Background: Self-efficacy has been proposed as a key predictor of alcohol treatment outcomes and a potential mechanism of success in achieving abstinence or drinking reductions following alcohol treatment. Integrative data analysis, where data from multiple studies are combined for analyses, can be used to synthesize analyses across multiple alcohol treatment trials by creating a commensurate measure and controlling for differential item functioning (DIF) to determine if alcohol treatments improve self-efficacy.

Method: The current study used moderated nonlinear factor analysis (MNLFA) to examine the effect of treatment on self-efficacy across four different treatment studies (N=3,720; 72.5% male, 68.4% non-Hispanic white). Self-efficacy was measured using the Alcohol Abstinence Self-Efficacy Scale (AASE) in the COMBINE Study (n=1,383) and Project MATCH (n=1,726), and the Drug Taking Confidence Questionnaire (DTCQ) in two studies of Telephone Continuing Care (TEL Study 1: n=303; TEL Study 2: n=212). DIF was examined across time, study, treatment condition, marital status, age, and sex.

Results: We identified 12 items from the AASE and DTCQ to create a commensurate measure of self-efficacy using MNLFA. Results indicated all active treatments, including cognitive-behavioral treatment, combined behavioral intervention, medication management, motivation enhancement treatment, telephone continuing care, twelve-step facilitation, and relapse prevention, were associated with significant increases in self-efficacy from baseline to posttreatment and these changes were maintained for up to a year. Importantly, treatment as usual in community settings, which consisted of weekly group therapy that included addiction counseling and twelve-step recovery support, was not associated with significant increases in self-efficacy.

Conclusions: Alcohol self-efficacy increases following treatment and numerous evidence-based treatments are associated with significant increases in self-efficacy, which are maintained over time. Community treatment that focuses solely on addiction counseling and twelve-step support may not promote increases in self-efficacy.

Corresponding Author: Katie Witkiewitz, Department of Psychology, University of New Mexico, MSC 03-2220, Albuquerque, NM 87131, katiew@unm.edu.

Introduction

Self-efficacy is defined as an individual's belief in their ability to organize and execute behaviors to achieve a desired outcome in prospective situations (Bandura, 1977). In the context of alcohol treatment, an individual with high self-efficacy is confident in their ability to abstain or reduce heavy drinking in potential alcohol use situations, such as social situations where alcohol is offered or in situations when an individual has previously used alcohol to cope with negative affect. On the other hand, an individual with low self-efficacy is unsure of his or her ability to abstain or reduce heavy drinking in these situations (Marlatt and Gordon, 1985). Previous research consistently finds higher self-efficacy is associated with better drinking outcomes following alcohol treatment (Kavanagh et al., 1996; Maisto et al., 2000; Sitharthan and Kavanagh, 1991), such that individuals with higher self-efficacy maintain abstinence longer (Vielva and Iraurgi, 2001), report fewer days of alcohol use (Brown et al., 2002), and show lower risk of returning to alcohol use after a period of abstinence (Allsop et al., 2000) than individuals with lower self-efficacy. Given these findings, self-efficacy may represent an important treatment mechanism leading to better alcohol treatment outcomes. The question remains whether specific treatments are more or less likely to increase self-efficacy.

Previous research has examined the differential effects of theoretically distinct behavioral treatments for alcohol use on self-efficacy. For example, much of the research examines differences in self-efficacy between cognitive-behavioral treatment and twelve-step oriented treatment. Theoretically, cognitive-behavioral treatment involves learning specific coping behaviors that help individuals navigate situations where the risk for alcohol use is high (Kadden et al., 1994), while twelve-step oriented treatment assumes that sustained abstinence is the only effective remedy to problematic alcohol use (Nowinski et al., 1995). These theoretical differences steer expectations for cognitive-behavioral therapy to exert greater effects on self-efficacy than twelve-step oriented treatment. However, two studies found no significant differences in self-efficacy between cognitive-behavioral treatment and twelve-step treatment (Glasner-Edwards et al., 2007; Johnson et al., 2006). Such findings are unexpected and would suggest that distinct behavioral treatments do not differ in their effects on self-efficacy.

A smaller but parallel line of research examines the differences in self-efficacy between behavioral treatments and medication. Theoretical discussions warn that medication may undermine an individuals' ability to build self-efficacy during treatment because an individual may attribute changes in alcohol use to medication rather than changes in their own behavior (Moncrieff and Drummond, 1997). Despite this warning, empirical research indicates the opposite may be true – medication produces greater changes in self-efficacy than behavioral treatments (Hartzler et al., 2011).

Collectively, these unexpected findings may be due to participants' differential response to alcohol treatment. Recent discussions in alcohol treatment precision medicine initiatives suggest specific subgroups of participants may benefit more from specific treatments than other subgroups of participants (e.g., Litten et al., 2012). Related to self-efficacy, there is evidence that distinct treatments may exert differential effects on self-efficacy based on simple demographic characteristics. For example, in a secondary analysis of Project MATCH (Project MATCH Research Group, 1997), Maisto and colleagues (2015) found that being male predicted greater increases in self-efficacy during treatment. However, no studies have comprehensively examined the effect of distinct alcohol treatments on self-efficacy, while adjusting for potential subgroup differences based on participant characteristics.

The lack of investigations into potential subgroups differences in measuring the construct of alcohol self-efficacy is problematic given that previous research has found differences in measures of general self-efficacy (Bonsaken et al., 2013; Peter et al., 2014) and domain-specific self-efficacy (Makransky et al., 2015; Riazi et al., 2014). Previous studies of self-efficacy in other domains have identified differences in item functioning by sex, age, time, and marital status, and that differential item functioning may also exist related to treatment received (e.g., receiving medication management vs CBT).

To examine the effect of treatment on self-efficacy across different treatment studies it is imperative to have similar measures of self-efficacy and outcomes in each study. Integrative data analysis (IDA) is an innovative framework for creating a cumulative science by accumulating knowledge from multiple studies via coordinated measurement or pooling samples and identifying and testing commensurate measures across studies (Curran and Hussong, 2009; Hofer and Piccinin, 2009; Hussong et al., 2013; Marcoulides and Grimm, 2017; McArdle et al., 2009; Shrout, 2009; Wilcox and Wang, 2021). Similar to meta-analysis, IDA extends findings beyond isolated clinical trials. IDA permits clinical research to generalize findings beyond a specific sample or study to maximize sample sizes, maximize sample heterogeneity, and examine differences across different studies and studied populations (Curran et al., 2014; Hofer and Piccinin, 2009; Wilcox and Wang, 2021). IDA also allows for larger sample sizes and permits testing hypotheses in specific subgroups (e.g., by sex and age groups) and to test effects of treatments that may differ by different subgroups (e.g., does treatment "X" have a greater effect on self-efficacy than treatment Y?) (Hussong et al., 2013). IDA can also be used to examine whether different study characteristics or participant characteristics within studies impact the measurement of constructs relevant to treatment outcomes (Hussong et al., 2013; Shrout, 2009; Wilcox and Wang, 2021).

IDA has increasingly been used in alcohol research to combine data from multiple studies, including investigations of alcohol use and sexual risk behavior (Walsh et al., 2017), changes in adolescent alcohol use over time (Silins et al., 2018), and brief alcohol interventions for college students (Mun et al., 2015). Given the lack of comprehensive understanding on alcohol treatments' effects on self-efficacy, the present study sought to use IDA to pool data across samples and examine the effect of various alcohol treatments on self-efficacy across studies and subgroups of participants (Curran and Hussong, 2009).

Method

Participants

Data used for this study were collected from four alcohol clinical trials ($N = 3,720$). These studies included the Project MATCH aftercare ($n = 774$) and outpatient samples ($n = 952$; Project MATCH research Group, 1997), the COMBINE study ($n = 1383$; COMBINE Study Group, 2003), and two studies of Telephone Continuing Care (TEL; TEL 1, $n = 359$; McKay et al., 2005; TEL2, $n = 252$; McKay et al., 2011). All four studies were randomized clinical trials. Participants in the total combined sample were, on average, 42.21 years old ($SD = 10.51$), predominantly male (72.5%), non-Hispanic White (68.4%) or African American (21.2%), and not married or cohabitating (66.2%).

The COMBINE study (COMBINE Study Group, 2003) randomized participants with alcohol dependence ($n = 1383$) from eleven research sites across the United States into nine treatment groups consisting of a combination of medical management (MM) or combined behavioral intervention (CBI) and medications (acamprosate, naltrexone, or placebo versions of each drug). Inclusion was determined by meeting the following criteria (1) alcohol dependence based on *the Diagnostic and Statistical Manual of Mental Health Disorders, Fourth Edition* (DSM-IV; American Psychiatric Association, 1994), (2) 4 to 21 days of abstinence; and (3) greater than 14 drinks (women) and 21 drinks (men) with at least 2 heavy drinking days--greater than or equal to 4 drinks (women) and 5 drinks (men) during a consecutive 30-day period within 90 days prior to baseline evaluation. Participants in COMBINE received treatment for 16 weeks (9 sessions of MM and up to 20 sessions of CBI). For evaluating predictor effects of study, the COMBINE study was used as the reference group.

Project MATCH (Project MATCH Research Group, 1997) randomized outpatients ($n = 952$) and aftercare patients (recruited from inpatient treatment; $n = 774$) with alcohol dependence or alcohol abuse from nine research sites across the United States into three treatment groups: cognitive behavioral therapy (CBT), motivation enhancement therapy (MET), or twelve-step facilitation (TSF). Eligibility was determined by meeting a diagnosis of alcohol dependence or abuse based on the *Diagnostic and Statistical Manual of Mental Health Disorders, Revised Third Edition* (DSM-III-R; American Psychiatric Association, 1987). Participants received up to 12 sessions of CBT and TSF and up to 4 sessions of MET over 12 weeks.

The TEL 1 study (McKay et al., 2005) recruited participants with cocaine and/or alcohol dependence ($n = 359$) after completing four weeks of intensive outpatient treatment in two outpatient treatment programs and randomized them to three continuing care treatments delivered over 12 weeks: weekly telephone based monitoring and brief counseling combined with weekly group support (TEL), weekly individual and group relapse prevention (RP), or treatment as usual (TAU, weekly group therapy that included addiction counseling and twelve-step recovery support). Participants were included if they had a previous diagnosis of cocaine or alcohol dependence (based on the DSM-IV criteria) at entrance to treatment and completed an intensive outpatient program (IOP) immediately prior.

The TEL 2 study (McKay et al., 2011) recruited participants with alcohol dependence (*n*=252) after completing three weeks of intensive outpatient treatment in two outpatient treatment programs and were randomized to three conditions: telephone-based monitoring for up to 18 months (TEL), up to 18 months of telephone-based monitoring and individual counseling using motivational interviewing and CBT sessions (TMC), or TAU for 3-4 months (once weekly group therapy that included addiction counseling and twelve-step recovery support). Participants were included if they had a previous diagnosis of alcohol dependence (based on the DSM-IV criteria) at entrance to treatment and completed an intensive outpatient program (IOP) immediately prior.

Measures

Alcohol Abstinence Self-Efficacy Scale (AASE; DiClemente et al., 1994).—The AASE is a self-report measure designed to assess the construct of self-efficacy as it applies to abstinence from alcohol. The measure asks participants to rate both perceived confidence and ability to abstain across 20 different situations on a 5-point Likert-type scale (1 = not at all confident, 5 = extremely confident). Participants in Project MATCH and COMBINE completed the AASE. In Project MATCH, the AASE was administered at baseline, post-treatment (3 months post-baseline), 6-month follow-up (9 months post-baseline), and 12-month follow-up (15 months post-baseline). In the COMBINE study, the AASE was administered at baseline, post-treatment (16 weeks post-baseline), and the 6-month follow-up (26 weeks, approximately 6.5 months, post-baseline). Internal consistency reliability of the AASE at baseline was $a = 0.95$ in MATCH and $a = 0.93$ in COMBINE. Reliability exceeded $a = 0.97$ at follow-up time points in MATCH and COMBINE.

Drug Taking Confidence Questionnaire (DTCQ; Annis et al., 1997).—The DTCQ is a 50-item self-report measure that ask participants to rate their perceived ability to resist the urge to drink heavily and cope in different situations. Participants rated their confidence on a 6-point Likert-type scale (0 = not at all confident and 100 = very confident). Participants in TEL1 and TEL2 completed the DTCQ at baseline, 3-months post-baseline, 6-months post-baseline, and 12-months post-baseline. The internal consistency reliability of the 50 item DTCQ exceeded $a = 0.97$ in TEL1 and TEL2 at all time points.

Data Analysis

Moderated nonlinear factor analysis (MNLFA) is one data analysis approach for pooling data in IDA (Curran and Hussong, 2009), in which indicators from different instruments across different studies can be combined. Further, MNLFA allows for testing differential item functioning (DIF) across potential covariates on the latent construct and individual indicators across multiple studies (Curran et al., 2014, 2018). All models described below were estimated with Mplus Version 8 (Muthén and Muthén, 2017) using maximum likelihood estimation.

Identification and Harmonization of Similar Items.—To perform the MNLFA participants were screened for missing data on demographic variables, those with missing data were excluded. The final sample size used in the MNLFA was 3,581. As outlined in the Supplementary Code and MNLFA input and output statements provided online at

https://osf.io/wa3tn, the first step in MNLFA was to identify similar items across the two self-efficacy measures. Similar items were identified and independently rated by two domain experts as items belonging to different measures that assessed similar content. We identified 12 self-efficacy items, see Table 1 for a description of items from the AASE and DTCQ. Since the AASE and DTCQ are scored on different Likert-type scales the DTCQ (scored from 0 to 100) was converted to the AASE scale (scored from 1 to 5): 0 (DTCQ) = 1 (AASE), 20 = 2, 40 = 2, 60 = 3, 80 = 4, and 100 = 5. The harmonization of Likert-type scales was based on distributions of the items in both samples. We also harmonized the time points across studies with baseline occurring before treatment in all studies and end of treatment occurring at 3 months post-baseline in the MATCH and TEL studies, and 4 months post-baseline in COMBINE.

Next, the data from the four RCTs were combined, which created a total of 20 treatment arms, therefore individual study arms were further combined based on similarity of interventional components into six treatment variables: Medical Management (MM; COMBINE MM arm), cognitive behavioral therapy (CBT; MATCH CBT and COMBINE CBI arms), motivational enhancement therapy (MET; MATCH MET arm), twelve-step facilitation (TSF; MATCH TSF), relapse prevention (RP; TEL1 RP arm), telephone based monitoring and counseling (TEL; TEL1 TEL arm; and TEL2 TEL and TMC arms), and treatment as usual (TAU; TEL1 and TEL2 TAU arms). All treatment variables were dummy coded with TAU as the reference.

Self-Efficacy Factor Structure.—Once the similar items and treatment variables were coded, the construct validity of the AASE, DTCQ, and self-efficacy items was evaluated by exploratory and confirmatory factor analysis (EFA and CFA, respectively) on a calibration sample. The calibration data set was created by randomly selecting one time-point for each participant. We elected to evaluate a parsimonious single factor model despite other studies using models with other factor configurations (DiClemente et al., 1994; Glozah et al., 2015, Kim et al., 2015; Ramo et al., 2009; Reilly et al., 1995; Sklar et al., 1997) to be consistent with previous studies of alcohol self-efficacy (Maisto et al., 2015; Mensinger et al., 2007; Witkiewitz et al., 2012). For each measure model fit was examined according *a priori* cutoff criteria (Comparative Fit Index (CFI) > 0.90; Tucker-Lewis Index (TLI) > 0.90; Root Mean Square Error of Approximation (RMSEA) < 0.08; and Standardized Root Mean Square Residual (SRMR) < 0.08; Hu and Bentler, 1999).

Moderated Nonlinear Factor Analysis.—Using the calibration data set, we first evaluated the effect of the covariates on the factor means and variances and then estimated 12 models, one model for each of the 12 similar items to test the effects of covariates on item intercepts and loadings. Specifically, to examine differential item functioning (DIF), we included covariates of time, study, treatment condition, marital status, age, and sex. These covariates were selected given prior studies of self-efficacy in other domains have identified differences in item functioning by sex, age, time, and marital status (Bonsaken et al., 2013; Makransky et al., 2015; Peter et al., 2014; Riazi et al., 2014) and because we were specifically interested in DIF by treatment received. Importantly, race was confounded with study, with nearly all African American participants recruited for the TEL studies, thus

we did not examine race/ethnicity in the current study. Next all significant covariate effects ($p < 0.10$) from examining (1) means and variances and (2) intercepts and loadings were included in a single model. From this single model an iterative model trimming process was conducted which consisted of removing non-significant covariates if they had a p-value > 0.05 and then rerunning the model. If a dummy coded covariate with multiple levels was significant (e.g., treatment) then all treatment dummy variables were included. This process continued until a final model was reached with only significant covariates remaining in the model.

Once the final model was estimated we derived expectation a posteriori (EAP) self-efficacy scores. The EAP is a predicted latent score of self-efficacy for each observation that are informed by covariate effects (Curran et al., 2018). Sensitivity of EAP scores were evaluated by correlating EAP self-efficacy values with values for the unadjusted raw self-efficacy commensurate items (derived from the AASE and DTCQ).

Self-Efficacy Latent Scores and Change in Self-Efficacy over Time.—After establishing a commensurate measure of alcohol self-efficacy, two separate analyses were conducted to examine (1) time and treatment related self-efficacy changes and (2) differences in self-efficacy across studies and treatments using the EAP scores. First, EAP self-efficacy means and variances were analyzed for linear trends over time and differences in self-efficacy between studies (COMBINE as reference) and across treatments (TAU as reference) were also examined. Mixed-effect models were used to analyze change in self-efficacy across all time points (collapsed across treatments) and between the baseline evaluation and end of treatment (3 months) for each treatment using EAP scores. Results of mixed effects models are reported as standardized mean difference (M /SD). Changes in unadjusted raw scores on the AASE, DTCQ, and raw items across time and within treatments were also analyzed and results are reported in supplemental tables.

Results

Identification of Similar Items

Twelve items from the DTCQ and AASE demonstrated significant construct overlap and were identified as similar items (see Table 1). Items between the AASE and DTCQ were matched based on theoretical overlap between items.

Self-Efficacy Factor Structure

Unidimensionality of the items was evaluated with exploratory factor analysis (EFA) using robust weighted least squares estimation with goemin rotation of all available data to examine dimensionality and item loadings and then subsequently with CFA using robust weighted least squares estimation. There was only one eigenvalue greater than 1.0 (first eigenvalue = 8.689, second eigenvalue = 0.718). Model fit of the one factor EFA model was not excellent, based on a significant χ^2 [χ^2 = 5520.62 (54), $p < 0.001$] and RMSEA = 0.19 (90% CI: 0.183, 0.192), however CFI = 0.953 and SRMR = 0.054 were acceptable and all item loadings exceeded 0.78. The one-factor CFA was also not excellent due to a significant

χ^2 [χ^2 = 5520.63 (54), $p < 0.001$] and a RMSEA = 0.19 (90% CI: 0.183, 0.192), however the CFI = 0.953 and SRMR = 0.042 were acceptable.

Moderated Nonlinear Factor Analysis

A unidimensional MNLFA was fit to the 12-item, 1-factor model of alcohol self-efficacy items, and included moderation by demographic variables (age, marital status, and sex), time, study membership, and treatment condition to test for the effects of these covariates on factor mean, factor variance, item intercepts, and item loadings. [1]

There were three items with significant DIF on item intercepts and/or loadings: "feeling depressed," "urge to try just one," and "social pressure." Item intercepts reflect the level of self-efficacy associated with endorsement of an item (see Table 2). Additionally, there was DIF on the loadings for study for the items "feeling depressed" (MATCH and TEL2), "urge to try just one drink" (MATCH and sex), and "social pressure" (MATCH and TEL1). The significant loadings indicate an interaction between study characteristics, measurement characteristics, or participant characteristics with levels of latent self-efficacy and thus probability of endorsement is non-uniform based on study characteristics and sex. Expectation a posteriori (EAP) scores for the alcohol self-efficacy latent factor were then created for each observation.

Self-Efficacy Latent Scores and Change in Self-Efficacy over Time

Only time, study, and treatment condition were significantly associated with the latent self-efficacy factor mean and variance. Latent self-efficacy mean and variance both increased significantly over time, mean: $B(SE) = 0.319(0.035)$, $p < 0.001$; variance $B(SE) = 0.346$ (0.033), $p < 0.001$. Individuals in the MATCH sample, had higher average self-efficacy than COMBINE, $B(SE) = 0.181(0.087)$, $p = 0.037$), and greater variance in self-efficacy, $B(SE) = 0.340(0.105)$, $p = 0.001$. Mean EAP self-efficacy scores were significantly higher in TEL, RP, and TAU treatment conditions, which is consistent with recruitment of individuals after already completing intensive outpatient treatment in the TEL1 and TEL2 studies. Specifically, compared to TAU, individuals in MM, CBT, MET, and TSF had lower average self-efficacy (MM: $B(SE) = -0.936(0.265)$, $p = 0.001$; CBT/CBI: $B(SE) = -0.898(0.262)$, $p = 0.001$; MET: $B(SE) = -0.930(0.279)$, $p = 0.001$; TSF: $B(SE) = -0.963(0.280)$, $p = 0.001$, and individuals in any TEL and RP conditions had significantly higher average self-efficacy compared to TAU, $B(SE) = 0.777(0.216)$, $p < 0.001$, and all other treatments tested in COMBINE and MATCH. Compared to TAU, the variance of self-efficacy was significantly smaller for the MM, CBT/CBI, and MET conditions, MM: $B(SE) = -0.369 (0.170)$, $p = 0.030$; CBT/CBI: $B(SE) = -0.390(0.165)$, $p = 0.018$; MET: $B(SE) = -0.391(0.188)$, $p = 0.038$, indicating less variability in the self-efficacy construct for those who received these treatments.

EAP Self-Efficacy Scores over Time.—As demonstrated in Table 3, the EAP self-efficacy scores demonstrated a significant increase from baseline to all other time points and

---

[1]Model fit of the MNLFA model was also estimated using weighted least squares estimation to obtain fit indices that are comparable to the CFA models. Results indicated acceptable model fit based on CFI = 0.98 and TLI = 0.98, and a better fitting model than the CFA results based on RMSEA = .095 (90% CI 0.094, 0.097).

did not significantly decrease across time points (all $ps > 0.21$). When comparing the effect of treatments from baseline to end of treatment (3 months), all treatments except TAU were associated with significant increases in self-efficacy (see Figure 1, Table 4).

Observed Self-Efficacy Scores over Time.—Analyses of scores on the observed (i.e., raw) self-efficacy items (i.e., summing the 12 items into a total self-efficacy score, unadjusted for covariate effects) produced somewhat different results than what was found using the EAP scores. First, the unadjusted observed self-efficacy scores indicated an initial significant increase in self-efficacy followed by a steady decline in observed self-efficacy scores between three and six ($p < 0.001$) and 12-months ($p < 0.001$), but no decline was observed in the EAP scores. Examining change in the unadjusted observed self-efficacy scores over time within each study by individual treatment conditions produced results largely consistent with the EAP self-efficacy scores with two exceptions: individuals assigned to the RP and TEL treatment conditions in the TEL1 study did not demonstrate a significant increase in observed self-efficacy scores (see Supplement Tables 3-6).

Discussion

In this study, we performed IDA by pooling data from four clinical trials to assess the effect of study treatments on alcohol self-efficacy. The active treatments, including cognitive-behavioral treatment, combined behavioral intervention, medication management, motivation enhancement treatment, telephone continuing care, twelve-step facilitation, and relapse prevention, were associated with a significant increase in self-efficacy from baseline to post-treatment and these changes were maintained up to one year. Importantly, treatment as usual in community settings, which consisted of weekly group therapy that included addiction counseling and twelve-step recovery support, was not associated with increased self-efficacy, above and beyond the levels of self-efficacy already achieved in intensive outpatient treatment.

Self-efficacy has been widely studied as one potential mechanism of change in alcohol treatment and has been shown to be a robust predictor of alcohol treatment outcomes (Allsop et al., 2000; Brown et al., 2002; Kavanagh et al., 1996; Maisto et al., 2000; Sitharthan and Kavanagh, 1991; Vielva and Iraurgi, 2001). Adamson and colleagues (2009) conducted a review of predictors of alcohol treatment outcome and found self-efficacy was the most consistent predictor. Additionally, similar to other studies examining treatment-related changes in self-efficacy over time (McKellar et al., 2008), increases in self-efficacy were maintained at follow-up. The current study demonstrated that across studies and treatment conditions, alcohol self-efficacy increased.

The effects of CBT, MET, RP, MM, TEL, and TSF on self-efficacy were robust with significant increases in self-efficacy from baseline to 3-months following treatment. These findings suggest that each of the specific treatments resulted in increases in self-efficacy from baseline to post-treatment. The fact that diverse treatments are effective in mobilizing self-efficacy could indicate that self-efficacy is a common target across treatments, or could indicate that changes in self-efficacy correspond to or are a byproduct of changes in behavior (i.e., drinking reductions achieved during treatment; Kadden and Litt, 2011).

The fact that medication management was also effective in increasing self-efficacy suggests the latter explanation may be more likely. An experimental study that attempted to causally manipulate self-efficacy among individuals who were motivated to quit smoking found those assigned to a high self-efficacy condition were more likely to quit smoking, but changes in self-reported self-efficacy did not mediate the effects of the intervention on smoking outcomes (Shadel et al., 2017). Interestingly, individual outcome expectancies did significantly mediate the effects of the self-efficacy intervention on smoking outcomes. Future work should continue to focus on whether self-efficacy is a mechanism of change, whether it is potentially a byproduct of behavior change (Kadden and Litt, 2011), and/or whether self-reported self-efficacy is independent from other cognitive factors that may explain behavior change.

Although there was a high degree of concordance between the EAP and unadjusted self-efficacy scores, there were some differences between EAP scores and unadjusted scores. Specifically, unadjusted self-efficacy scores demonstrated a significant decrease in self-efficacy from immediate post-treatment to subsequent follow-up measurements. These findings would suggest that unadjusted self-efficacy scores decrease following treatment, however the analyses with the EAP scores do not indicate the erosion of self-efficacy across time following treatment. Similarly, the effect of RP and TEL on the change in self-efficacy from baseline to post-treatment was not significant in the TEL1 study using the unadjusted self-efficacy score, but did produce a significant effect using the EAP scores. Thus, findings diverged for the effect of RP and TEL on change in self-efficacy. Prior simulation work has found that raw scores may lead to biased inferential tests due to covariate effects that are not estimated, whereas EAP scores seem to mitigate these biases (Curran et al., 2018). Three of the 12 of the items used to develop the EAP self-efficacy scores demonstrated differential item functioning by gender and study. Generally, these findings suggest that certain situations are more applicable for self-efficacy for different study populations and in different treatment groups.

Limitations

The current study utilized data collected from large and highly rigorous alcohol clinical trials, which is a strength and also a limitation. Combining treatments across studies prevents us from considering the randomized effective treatment. In other words, even though the individuals in each trial were randomized, the combination of individuals across treatments were not randomized. The current study was more focused on within-person changes in self-efficacy following treatment, however we cannot say that one treatment was better in a way that capitalizes on the randomized clinical trial design methodology used within each study.

All four clinical trials had exclusion criteria that limit generalizability of the current findings. The samples were also predominantly male and had a large representation of white participants, which limited our ability to assess for subgroup differences by race and ethnicity. It is also important to note that study and race were confounded, such that participants in the TEL studies were predominantly Black/African American. This study level difference was not explicitly captured in the analyses, which is a significant limitation.

Author Manuscript

Similarly, the AASE used in the MATCH and COMBINE studies assessed abstinence self-efficacy, whereas the DTCQ used in the TEL studies assessed self-efficacy to resist urges to drink heavily. It is unclear if higher scores in the TEL studies may reflect greater self-efficacy to resist heavy drinking or other study level differences. The poor model fit, based on RMSEA and significant chi-square, is also a limitation. We proceeded with testing the one factor model given prior work has used a total score (assuming a one factor model) and given complexities in testing MNLFA models with multiple factors, however future work could consider examining a multidimensional structure based on the subscales of the AASE and DTCQ to examine whether subscales are invariant across participant characteristics and time. Additionally, future work may consider testing the models using a multilevel IDA framework (Shrout, 2009; Wilcox and Wang, 2021).

Comparing the results of harmonized EAP scores with raw scores that did not adjust for covariates and time, leads to different conclusions in some cases. This is most evident for the effect of time, where the unadjusted raw self-efficacy scores demonstrate a significant decrease from immediate post-treatment to subsequent follow-up measurements. These findings would suggest that unadjusted self-efficacy scores decrease following treatment, however the analyses with the EAP scores across time following treatment show a more optimistic pattern of stability of self-efficacy following treatment. These findings should be considered preliminary and require replication, given the small number of studies included and the lack of multiple common items across studies.

Conclusions

The current study provides novel information about the measurement of self-efficacy in alcohol clinical trials. Future work should replicate these findings, and include other demographic factors, particularly race and ethnicity. Research should continue to examine the measurement of constructs that are central to alcohol treatment outcomes research, such as affect, craving, executive function, and social support. Examining measures across patient populations and in different settings and contexts has the potential to improve alcohol use disorder treatment and precision medicine approaches to matching patient populations to specific alcohol use disorder treatment protocols (Kranzler and McKay, 2012).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding:

References

Adamson SJ, Sellman JD, Frampton CM (2009) Patient predictors of alcohol treatment outcome: a systematic review. J Subst Abuse Treat 36:75–86. [PubMed: 18657940]

Allsop S, Saunders B, Phillips M (2000) The process of relapse in severely dependent male problem drinkers. Addiction 95:95–106. [PubMed: 10723834]

American Psychiatric Association (1987) Diagnostic and statistical manual of mental disorders. 3rd ed. American Psychiatric Association, Washington, DC.

American Psychiatric Association (1994) Diagnostic and statistical manual of mental disorders. 4th ed. American Psychiatric Association, Washington, DC.

Annis HM, Sklar SM, Turner NE (1997) The Drug-Taking Confidence Questionnaire (DTCQ): User's Guide. Toronto, Canada: Addiction Research Foundation, Centre for Addiction and Mental Health.

Bandura A (1977) Self-efficacy: Toward a unifying theory of behavioral change. Psychol Rev 84:191–215. [PubMed: 847061]

Bonsaksen T, Kottorp A, Gay C, Fagermoen MS, Lerdal A (2013) Rasch analysis of the General Self-Efficacy Scale in a sample of persons with morbid obesity. Health Qual Life Outcomes 11:202. [PubMed: 24268204]

Brown TG, Seraganian P, Tremblay J, Annis H (2002) Process and outcome changes with relapse prevention versus 12-Step aftercare programs for substance abusers. Addiction 97:677–689. [PubMed: 12084137]

COMBINE Study Research Group (2003) Testing combined pharmacotherapies and behavioral interventions in alcohol dependence: rationale and methods. Alcohol Clin Exp Res 27:1107–1122. [PubMed: 12878917]

Curran PJ, Cole VT, Bauer DJ, Rothenberg WA, Hussong AM (2018) Recovering predictor-criterion relations using covariate-informed factor score estimates. Struct Equ Modeling 25:860–875. [PubMed: 31223223]

Curran PJ, Hussong AM (2009) Integrative data analysis: the simultaneous analysis of multiple data sets. Psychol Methods 14:81–100. [PubMed: 19485623]

Curran PJ, McGinley JS, Bauer DJ, Hussong AM, Burns A, Chassin L, Sher K, Zucker R (2014) A moderated nonlinear factor model for the development of commensurate measures in integrative data analysis. Multivariate Behav Res 49:214–231. [PubMed: 25960575]

DiClemente CC, Carbonari JP, Montgomery RP, Hughes SO (1994) The alcohol abstinence self-efficacy scale. J Stud Alcohol 55:141–148. [PubMed: 8189734]

Glasner-Edwards S, Tate SR, McQuaid JR, Cummins K, Granholm E, Brown SA (2007) Mechanisms of action in integrated cognitive-behavioral treatment versus twelve-step facilitation for substance-dependent adults with comorbid major depression. J Stud Alcohol Drugs 68:663–672. [PubMed: 17690799]

Glozah FN, Adu NAT, Komesuor J (2015) Assessing alcohol abstinence self-efficacy in undergraduate students: psychometric evaluation of the alcohol abstinence self-efficacy scale. Health Qual Life Outcomes 13:189. [PubMed: 26607755]

Hartzler B, Witkiewitz K, Villarroel N, Donovan D (2011) Self-efficacy change as a mediator of associations between therapeutic bond and one-year outcomes in treatments for alcohol dependence. Psychol Addict Behav 25:269–278. [PubMed: 21443293]

Hofer SM, Piccinin AM (2009) Integrative data analysis through coordination of measurement and analysis protocol across independent longitudinal studies. Psychol Methods 14:150–164. [PubMed: 19485626]

Hu L, Bentler PM (1999) Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Model 6:1–55.

Hussong A, Curran PJ, Bauer DJ (2013) Integrative data analysis in clinical psychology research. Annu Rev Clin Psychol 9:61–89. [PubMed: 23394226]

Johnson JE, Finney JW, Moos RH (2006) End-of-treatment outcomes in cognitive-behavioral treatment and 12-step substance use treatment programs: do they differ and do they predict 1-year outcomes? J Subst Abuse Treat 31:41–50. [PubMed: 16814009]

Kadden RM, Carroll KM, Donovan DM, Cooney N, Monti P, Abrams D, Litt M, Hester R (1995) Cognitive-behavioral coping skills therapy manual. National Institute on Alcohol Abuse and Alcoholism, Rockville, MD.

Kadden RM, Litt MD (2011) The role of self-efficacy in the treatment of substance use disorders. Addict Behav 3612:1120–1126.

Kavanagh DJ, Sitharthan T, Sayer GP (1996) Prediction of results from correspondence treatment for controlled drinking. Addiction 91:1539–1545. [PubMed: 8917921]

Kim SY, Ahn S, Lim TS (2015) Predicting alcohol misuse among college students in the US and South Korea. J Ethn Subst Abuse 14:308–322. [PubMed: 26161809]

Kranzler HR, McKay JR (2012) Personalized treatment of alcohol dependence. Curr Psychiatry Rep 14:486–493. [PubMed: 22810115]

Litten RZ, Egli M, Heilig M, Cui C, Fertig JB, Ryan ML, Falk DE, Moss H, Huebner R, Noronha A (2012) Medications development to treat alcohol dependence: a vision for the next decade. Addict Biol 17:513–527. [PubMed: 22458728]

Maisto SA, Connors GJ, Zywiak WH 2000. Alcohol treatment changes in coping skills, self-efficacy, and levels of alcohol use and related problems 1 year following treatment initiation. Psychol Addict Behav 14:257–266. [PubMed: 10998951]

Maisto SA, Roos CR, O'Sickey AJ, Kirouac M, Connors GJ, Tonigan JS, Witkiewitz K (2015) The indirect effect of the therapeutic alliance and alcohol abstinence self-efficacy on alcohol use and alcohol-related problems in Project MATCH. Alcohol Clin Exp Res 393:504–513.

Makransky G, Rogers ME, Creed PA (2015) Analysis of the construct validity and measurement invariance of the career decision self-efficacy scale: a Rasch model approach. J Career Assess 23:645–660.

Marcoulides KM, Grimm KJ (2017). Data integration approaches to longitudinal growth modeling. Educ Psychol Meas 77: 971–989. [PubMed: 29795941]

Marlatt GA, Gordon JR. (1985) Relapse prevention: maintenance strategies in addictive behavior change. Guilford Press, New York.

McCardle JJ, Grimm K, Hamagami F, Bowles R, Meredith W (2009). Modeling life-span growth curves of cognition using longitudinal data with multiple samples and changing scales of measurement. Psychol Methods 14: 126–149. [PubMed: 19485625]

McKay JR, Lynch KG, Shepard DS, Pettinati HM (2005) The effectiveness of telephone-based continuing care for alcohol and cocaine dependence: 24-month outcomes. Arch Gen Psychiatry 622:199–207.

McKay JR, Van Horn D, Oslin DW, Ivey M, Drapkin ML, Coviello DM, Yu Q, Lynch KG (2011) Extended telephone-based continuing care for alcohol dependence: 24-month outcomes and subgroup analyses. Addiction 106:1760–1769. [PubMed: 21545667]

McKellar J, Ilgen M, Moos BS, Moos R (2008) Predictors of changes in alcohol-related self-efficacy over 16 years. J Subst Abuse Treat 35:148–155. [PubMed: 18037604]

Mensinger JL, Lynch KG, TenHave TR, McKay JR (2007) Mediators of telephone-based continuing care for alcohol and cocaine dependence. J Consult Clin Psychol 75:775–784. [PubMed: 17907859]

Moncrieff J, Drummond DC (1997) New drug treatments for alcohol problems: a critical appraisal. Addiction 92:939–947. [PubMed: 9376777]

Mun EY, de la Torre J, Atkins DC, White HR, Ray AE, Kim SY, Jiao Y, Clarke N, Huo Y, Larimer ME, Huh D, The Project INTEGRATE Team. (2015). Project INTEGRATE: An integrative study of brief alcohol interventions for college students. Psychol Addict Behav, 29:34–48. [PubMed: 25546144]

Muthén LK, Muthén BO. (1998-2017). Mplus User's Guide. Eigth Edition. Los Angeles, CA: Muthén & Muthén.

Nowinski J, Baker S, Carroll KM (1995) Twelve step facilitation therapy manual. National Institute on Alcohol Abuse and Alcoholism, Rockville, MD.

Peter C, Cieza A, Geyh S (2014) Rasch analysis of the General Self-Efficacy Scale in spinal cord injury. J Health Psychol 19:544–555. [PubMed: 23463793]

Pilkonis PA, Yu L, Dodds NE, Johnston KL, Lawrence SM, Daley DC (2016) Validation of the alcohol use item banks from the Patient-Reported Outcomes Measurement Information System (PROMIS). Drug Alcohol Depend 161:316–322. [PubMed: 26936412]

Project MATCH Research Group (1997) Matching alcoholism treatments to client heterogeneity. J Stud Alcohol 58:7–29. [PubMed: 8979210]

Ramo DE, Myers MG, Brown SA (2009) Psychometric properties of a revised form of the Drug-taking Confidence Questionnaire for use with adolescents. J Child Adolesc Subst Abuse 18:24–42 [PubMed: 20354581]

Reilly PM, Sees KL, Shopshire MS, Hall SM, Delucchi KL, Tusel DJ, Banys P, Clark HW, Piotrowski NA (1995) Self-efficacy and illicit opioid use in a 180-day methadone detoxification treatment. J Consult Clin Psychol 63:158–162. [PubMed: 7896984]

Riazi A, Aspden T, Jones F (2014) Stroke Self-efficacy Questionnaire: a Rasch-refined measure of confidence post stroke. J Rehab Med 46:406–412.

Shadel WG, Martino SC, Setodji C, Cervone D, Witkiewitz K (2017) Does self-efficacy causally influence initial smoking cessation? An experimental study. Addict Behav 73:199–203. [PubMed: 28549331]

Shrout P (2009). Short and long views of integrative data analysis: Comments on conributions to the special issue. Psychol Methods 14:177–181. [PubMed: 19485628]

Silins E, Horwood LJ, Najman JM, Patton GC, Toumbourou JW, Olsson CA, Hutchinson DM, Degenhardt L, Fergusson D, Becker D, Boden JM, Borschmann R, Plotnikova M, Youssef GJ, Tait RJ, Clare P, Hall WD, Mattick RP; Cannabis Cohorts Research Consortium. (2018). Adverse adult consequences of different alcohol use patterns in adolescence: an integrative analysis of data to age 30 years from four Australasian cohorts. Addiction. 113:1811–1825 [PubMed: 29749666]

Sitharthan T, Kavanagh DJ (1991) Role of self-efficacy in predicting outcomes from a programme for controlled drinking. Drug Alcohol Depend 27:87–94. [PubMed: 2029863]

Sklar SM, Annis HM Turner NE (1997) Development and validation of the drug-taking confidence questionnaire: A measure of coping self-efficacy. Addict Behav 22:655–670. [PubMed: 9347068]

Vielva I, Iraurgi I (2001) Cognitive and behavioural factors as predictors of abstinence following treatment for alcohol dependence. Addiction 96:297–303. [PubMed: 11182875]

Walsh JL, Weinhardt LS, Kalichman SC, Carey MP.(2017) Using Integrative Data Analysis to Examine Changes in Alcohol Use and Changes in Sexual Risk Behavior Across Four Samples of STI Clinic Patients. Ann Behav Med. 51(1):39–56. [PubMed: 27550626]

Wilcox KT, Wang L (2021, 7 29). Modeling approaches for cross-sectional integrative data analysis: Evaluations and recommendations. Psychol Methods. Advance online publication. 10.1037/met0000397

Witkiewitz K, Donovan DM, Hartzler B (2012) Drink refusal training as part of a combined behavioral intervention: effectiveness and mechanisms of change. J Consult Clin Psychol 803:440–449.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript
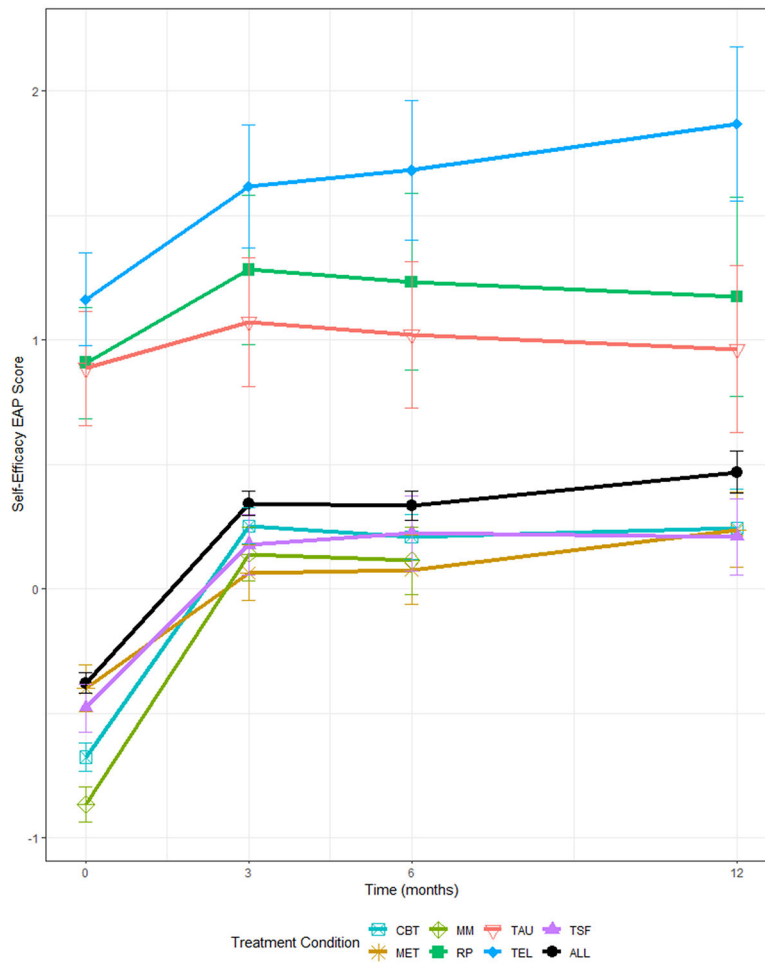
Figure 1. Change in Self-Efficacy Over Time

Note: Error bars denote 95% confidence intervals; CBT = Cognitive Behavioral Therapy, MM = Medication Management, TAU = Treatment as Usual, TSF = Twelve Step Facilitation, MET = Motivation Enhancement Therapy, RP = Relapse Prevention, and TEL = Telehealth, & ALL = grand mean across all individuals and conditions.

Table 1.

Description of Items that Were Rated to be Similar

	AASE	DTCQ
Root	Confident not to drink alcohol…	I would be able to resist the urge to use alcohol…
Response Levels and Anchors	1 = Not at all, 2 = Not very, 3 = Moderately, 4 = Very, 5 = Extremely	0 = not at all confident, 20, 40, 60, 80, 100 = very confident (numbers indicate percent confidence)
Similar item stems		
Depression	3. When I am feeling depressed	1. If I were depressed about things in general
Headache	13. When I have a headache	42. If I had a headache or was in physical pain
Worry/Anxiety	6. When I am worried	12. If I felt anxious or tense about something
Urge to Try Just One	7. When I have the urge to try just one drink to see what happens	25. If I wanted to find out whether I could take a drink
Will Power	10. When I want to test my will power over drinking	15. If I wanted to prove to myself that alcohol was not a problem for me
Anger	18. When I was feeling angry inside	41. If I were angry at the way things had turned out
Social Pressure	17. When people encourage me to drink	38. If I were pressured to drink and felt that I couldn't refuse
See Others Drinking	15. When I see others drinking	48. If I were with a group of people and everyone was drinking
Hopelessness	16. When I have the sense that everything is going wrong	4. If I felt there was nowhere left to turn
Excitement/Celebration	20. When I am excited/celebrating	23. If I felt excited about something
Withdrawal Symptoms	1. In agony from stopping/withdrawing	2. If I felt shaky, sick, or nauseous
Sleep/Tiredness	12. When I am physically tired	5. If I would have trouble sleeping

Note. AASE = Alcohol Abstinence Self-Efficacy Scale; DTCQ = Drug Taking Confidence Questionnaire.

Table 2.

Results from final MNLFA Alcohol Self-Efficacy Model with Significant Covariate Effects on Item Intercepts and Item Loadings

Item	Covariate	Intercept (SE)	*p*-value	Loading (SE)	*p*-value
Self-efficacy item 1 – "Depressed"	MATCH	−0.067 (0.112)	0.547	−0.253 (0.095)	0.008
	TEL1	−0.065 (0.180)	0.716	−0.154 (0.150)	0.305
	TEL2	−0.213 (0.236)	0.366	−0.476 (0.168)	0.005
Self-efficacy item 4 – "Urge to try just one drink"	MATCH	0.557 (0.101)	<0.001	−0.212 (0.087)	0.015
	TEL1	−0.029 (0.169)	0.864	−0.208 (0.131)	0.112
	TEL2	−0.213 (0.213)	0.318	−0.254 (0.153)	0.096
	Sex (male=1)	−0.086 (0.098)	0.380	0.203 (0.076)	0.008
Self-efficacy item 7 – "Social pressure"	MATCH	0.303 (0.108)	0.005	−0.243 (0.093)	0.009
	TEL1	−0.107 (0.179)	0.551	−0.285 (0.140)	0.042
	TEL2	−0.153 (0.213)	0.473	−0.160 (0.178)	0.368

Note. COMBINE is the reference group for dummy coded study contrasts.

Table 3.

Alcohol Self-Efficacy EAP Scores at Each Time Point

| Time | Alcohol Self-Efficacy EAP | | | |
	Baseline	3 mo.	6 mo.	12 mo.
Baseline	[−0.331]	<.001	<.001	<.001
3 mo.	−0.456	[0.125]	<.001	0.214
6 mo.	−0.454	0.002	[0.123]	0.189
12 mo.	−0.483	−0.027	−0.029	[0.152]

Note. EAP = Expectation a posteriori scores from moderated nonlinear factor analysis; bracketed diagonal = standardized value at each time point; below diagonal is the standardized difference between time-points; above diagonal is the *p*-value.

Table 4.

Alcohol Self-Efficacy EAP Scores From Start to End Of Treatment Using MNLFA to Pool Studies and Treatment Conditions

	Alcohol Self-Efficacy EAP			
	Std. Difference	Standard Error	Lower CI	Upper CI	p-value
CBT	0.590	0.027	0.537	0.642	<0.001
MET	0.298	0.041	0.218	0.377	<0.001
MM	0.643	0.041	0.562	0.723	<0.001
RP	0.243	0.087	0.072	0.414	0.005
TAU	0.119	0.075	−0.027	0.265	0.111
TEL	0.272	0.067	0.141	0.402	<0.001
TSF	0.411	0.041	0.331	0.490	<0.001

Note: EAP = expectation a posteriori scores from moderated nonlinear factor analysis ; CBT = cognitive behavioral therapy; MET = motivational enhancement therapy; MM = medical management; RP = relapse prevention; TAU = treatment as usual; TEL = telehealth; TSF = twelve-step facilitation.