



A Structural Equation Modeling Approach to Meta-analytic Mediation Analysis Using Individual Participant Data: Testing Protective Behavioral Strategies as a Mediator of Brief Motivational Intervention Effects on Alcohol-Related Problems

David Huh¹ · Xiaoyin Li² · Zhengyang Zhou³ · Scott T. Walters² · Scott A. Baldwin⁴ · Zhengqi Tan³ · Mary E. Larimer⁵ · Eun-Young Mun²

Accepted: 1 October 2021 / Published online: 12 November 2021
© The Author(s) 2021

Abstract

This paper introduces a meta-analytic mediation analysis approach for individual participant data (IPD) from multiple studies. Mediation analysis evaluates whether the effectiveness of an intervention on health outcomes occurs because of change in a key behavior targeted by the intervention. However, individual trials are often statistically underpowered to test mediation hypotheses. Existing approaches for evaluating mediation in the meta-analytic context are limited by their reliance on aggregate data; thus, findings may be confounded with study-level differences unrelated to the pathway of interest. To overcome the limitations of existing meta-analytic mediation approaches, we used a one-stage estimation approach using structural equation modeling (SEM) to combine IPD from multiple studies for mediation analysis. This approach (1) accounts for the clustering of participants within studies, (2) accommodates missing data via multiple imputation, and (3) allows valid inferences about the indirect (i.e., mediated) effects via bootstrapped confidence intervals. We used data ($N = 3691$ from 10 studies) from Project INTEGRATE (Mun et al. *Psychology of Addictive Behaviors*, 29, 34–48, 2015) to illustrate the SEM approach to meta-analytic mediation analysis by testing whether improvements in the use of protective behavioral strategies mediate the effectiveness of brief motivational interventions for alcohol-related problems among college students. To facilitate the application of the methodology, we provide annotated computer code in R and data for replication. At a substantive level, stand-alone personalized feedback interventions reduced alcohol-related problems via greater use of protective behavioral strategies; however, the net-mediated effect across strategies was small in size, on average.

Keywords Complex synthesis · Integrative data analysis · Indirect effect · Bootstrap inference with multiple imputation · Brief alcohol intervention

Introduction

Mediation analysis is used to evaluate whether the effects of an intervention on health outcomes occur because of change in a key behavior targeted by the intervention. Most of the existing methodological research and applications of mediation analysis have focused on individual studies. However, beyond assessing the overall effectiveness of a treatment, single-study intervention trials are frequently underpowered to evaluate pathways of change (Fritz et al., 2015). A meta-analytic approach to mediation analysis that leverages data from multiple studies provides an opportunity to test pathways of change with greater statistical power. However, the literature showing how to conduct mediation analysis in

✉ David Huh
dhuh@uw.edu

¹ School of Social Work, University of Washington, 4101 15th Ave NE, Box 354900, Seattle, WA 98105-6299, USA

² Department of Health Behavior and Health Systems, University of North Texas Health Science Center, Fort Worth, TX, USA

³ Department of Biostatistics and Epidemiology, University of North Texas Health Science Center, Fort Worth, TX, USA

⁴ Department of Psychology, Brigham Young University, Provo, UT, USA

⁵ Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA, USA

a meta-analytic context has been limited to aggregate data (Cheung & Chan, 2005). This paper focuses on methods for conducting mediation analysis using individual participant data (IPD) from multiple studies.

The most widely used method for combining data from multiple studies is meta-analysis using study-level, aggregate data (e.g., means, *SDs*, correlations); however, standard meta-analysis methods either do not lend themselves to mediation testing or do not accommodate IPD from multiple studies. For example, meta-regression is used to examine moderators of intervention effects—study-level predictors that are associated with the size of the effect—not mediation. In contrast, newer approaches using aggregate data, such as meta-analytic structural equation modeling (MASEM), can provide a test of mediation (i.e., indirect effects) when pooling data from multiple studies (Cheung, 2014, 2015). Correlation-based MASEM is a prevailing approach to meta-analytic mediation analysis in which correlation or covariance matrices extracted from published reports or generated from the raw data (Cheung & Chan, 2005, 2009) are combined to create a pooled correlation or covariance matrix that is subsequently analyzed using structural equation modeling (SEM; e.g., Wilson et al., 2016). Effect sizes and standard errors may also be utilized to test mediated effects via marginal likelihood synthesis, sequential Bayesian methods, or parameter-based MASEM (see van Zundert & Miočević, 2020 for a comparison).

However, because prevailing approaches for meta-analytic mediation analysis typically rely on aggregate data extracted from published reports, the findings may be confounded with study-level differences that are unrelated to the mechanism of interest. For example, Riley et al. (2010) illustrated a meta-regression of ten clinical trials for hypertension where the estimated treatment effect was smaller in men compared to women, whereas a one-step IPD meta-analysis that examined participant-level information directly within studies did not support a clinically significant difference in treatment effect by sex. The apparent superiority of treatment with women was an artifact of studies with larger proportions of female participants tending towards larger effect sizes, though for reasons unrelated to sex. Specifically, when treatment effects by sex were evaluated within studies, the differences in treatment response were not clinically significant. Consequently, the study-level summaries that are commonly utilized can make these approaches more prone to ecological inference bias. An advantage of MASEM is that within-study variables (e.g., sex) can be included in the model, which can avoid introducing ecological biases when individual-level data are aggregated and analyzed as study-level data (e.g., proportion of females in the study); however, this generally requires access to raw IPD.

A limitation of correlations as the input data for a mediation analysis is the loss of scale-level information since

correlation coefficients are standardized within each study to have a mean of zero and a standard deviation of one. This allows the pooling and comparison of the correlations across studies but assumes that the bivariate correlations correspond with the same range of values on the variable scales across intervention groups and levels of the outcome and mediator variables within studies. In practice, it is difficult to know whether these assumptions are reasonable without verifying them with IPD. If these assumptions are not met, then the resulting inference could be biased.

Furthermore, MASEM and existing approaches utilizing aggregate data are generally limited by the information disclosed in intervention reports, which frequently do not include all outcomes that were assessed (see Mun et al., 2021), let alone correlations among key variables of interest. Thus, MASEM and other mediation modeling approaches that rely on aggregate data may not be possible in many cases without access to IPD or unreported aggregate data. Finally, with only aggregate data, it is impossible to check and verify whether the original data were appropriately analyzed and reported (e.g., the assumptions of multivariate normal distribution, data that is missing at random).

Meta-analysis using IPD provides an opportunity to more rigorously evaluate the pathways by which treatments improve health outcomes at the individual level. Furthermore, a mediation analysis with IPD permits a longitudinal analysis that controls for baseline levels of (a) the mediator, (b) the outcome, and (c) any relevant covariates.

The current paper proposes an SEM approach using IPD that (a) accounts for the clustering of participants within studies, (b) accommodates missing data via multiple imputation, and (c) allows valid inferences about the indirect effect (i.e., mediated effect) via bootstrapped confidence intervals in an integrative data analysis (IDA) that estimates the entire model in one step, after previously establishing commensurate measures (see Hussong et al., 2013 for typical considerations for IDA). In this article, we first introduce the motivating research question and example data. Second, we outline a meta-analytic mediation modeling approach that can accommodate the clustered data structure of participants nested within studies. Third, we discuss how to estimate confidence intervals for the indirect and total effects of intervention for the purpose of statistical inference. Finally, we illustrate the meta-analytic mediation analysis using data drawn from Project INTEGRATE (Mun et al., 2015) and discuss the implications of our method for both methodological and substantive research.

The motivating research question is whether improvements in protective behavioral strategies (PBS) mediate the effectiveness of brief motivational interventions for alcohol-related problems among college students who drink. PBS are specific cognitive-behavioral strategies that can be used prior to or during alcohol consumption to reduce alcohol-related

problems (Martens et al., 2013). In the past two decades, promoting the use of PBS has become a common component of interventions for reducing alcohol-related problems among college drinkers (Ray et al., 2014). However, there has been mixed evidence on the extent to which improvements in PBS can explain the effect of brief motivational interventions on reducing alcohol use and related problems, with most evidence coming from cross-sectional data (Reid & Carey, 2015). We detail a longitudinal mediation analysis approach to evaluate whether improvements in PBS following brief motivational intervention are associated with subsequent reductions in alcohol-related problems among college students who drink.

Motivating Data: The Project INTEGRATE Study

The motivating data are drawn from Project INTEGRATE, a large-scale IPD meta-analysis project evaluating brief motivational interventions for college drinking across 24 independent intervention studies (Mun et al., 2015). From the Project INTEGRATE data set, we selected ten studies that were randomized controlled trials assessing PBS and alcohol-related problems at baseline and at least one post-baseline assessment. Participants in the included studies were randomized to a control group or one of three brief motivational interventions: (1) individually delivered motivational interviewing with personalized feedback (MI + PF), (2) stand-alone personalized feedback (PF), or (3) group-based motivational interviewing (GMI). Because PBS is not applicable for non-drinkers, we only included participants within each study who reported at least one drink in the

past 1 or 3 months, depending on the study, at post-baseline assessment. Table 1 summarizes the intervention arms and corresponding sample sizes for the combined sample of drinkers from the ten studies that met the study inclusion criteria. Eight of the 10 studies were two-arm trials that evaluated a single brief motivational intervention, whereas studies 9 and 21 evaluated two or more intervention groups.

The mediator variable, PBS, was measured using five different scales across the original studies, which were subsequently harmonized and made commensurate by using a generalized partial credit model (Muraki, 1992), which is an extension of the hierarchical two-parameter logistic item response theory (2-PL IRT) model that we reported for alcohol-related problems (Huo et al., 2015). The measurement work to establish PBS trait scores can be found in Mun et al. (2015, 2016). With respect to the motivating data, studies 2, 8a, 8b, 8c, and 9 used the 10-item Protective Behavioral Strategies (PBS; American College Health Association, 2001) measure; studies 16, 18, and 21 used the 15-item Protective Behavioral Strategies Scale (PBSS; Martens et al., 2005); and studies 12 and 22 used the seven-item Drinking Restraining Strategies (DRS; Wood et al., 2007) measure. Study 22 incorporated an additional nine-item measure asking about Drinking Strategies. These scales shared similarly worded items, from which five collapsed items across scales provided overlap across studies when estimating item parameters.

The outcome variable, alcohol-related problems, was assessed using six different scales across the original studies. We used latent trait scale scores estimated from hierarchical, 2-PL IRT models for multiple groups to establish commensurate alcohol-related problems trait scores for all participants across studies and time (Huo et al., 2015; Mun et al., 2015).

Table 1 The combined sample by intervention group and study

Study	Intervention group (<i>n</i>)				Follow-up (in months)	Reference
	Control	MI + PF	PF	GMI		
2	74	–	70	–	2	White et al. (2008)
8a	429	–	398	–	12	Larimer et al. (2007)
8b	585	–	544	–	12	Larimer et al. (2007)
8c	131	–	113	–	12	Larimer et al. (2007)
9	78	79	77	80	3	Lee et al. (2009)
12	81	76	–	–	1	Wood et al. (2007)
16	86	–	–	97	1	LaBrie et al. (2009)
18	67	–	73	–	1	Martens et al. (2010)
21	66	68	59	–	3	Walters et al. (2009)
22	189	171	–	–	12	Wood et al. (2010)
All	1786	394	1334	177		

The follow-up (in months) is the first post-baseline assessment for which both mediation and outcome data were collected in the study.

MI + PF individually delivered motivational interviewing intervention with personalized feedback, *PF* stand-alone personalized feedback intervention, *GMI* group motivational interviewing intervention

With respect to the motivating data, studies 2, 8a, 8b, 8c, 9, 16, and 21 used the Rutgers Alcohol Problem Index (RAPI; White & Labouvie, 1989); studies 8a, 8b, 8c, 9, 12, 16, and 22 used the Young Adult Alcohol Problems Screening Test (YAAPST; Hurlbut & Sher, 1992); study 12 also used the Alcohol Dependence Scale (Skinner & Allen, 1982; Skinner & Horn, 1984); study 18 used the Brief Young Adult Alcohol Consequences Questionnaire (BYAACQ; Kahler et al., 2005); and study 21 used the Alcohol Use Disorders Identification Test (AUDIT; Saunders et al., 1993). For readers interested in the technical details regarding how the measures of PBS and alcohol problems used in the motivating data were made commensurate, the harmonization work is discussed extensively in earlier reports (Huo et al., 2015; Mun et al., 2015, 2016, 2019).

The sample for the present analysis included a total of 3691 students, with approximately two-thirds (63.8%) female. Most of the students identified as White (78.3%), and just over half of the participants (56.2%) were first-year or incoming college students. Table 2 provides a descriptive summary of all variables, including rates of missing data, by study and time point.

Meta-analytic Mediation Model for Pretest–Posttest Designs

Clinical trials commonly use pretest–posttest designs in which participants are assessed at baseline and one or more follow-ups. In the current motivating data, half of the studies included a single follow-up within 12 months post-intervention (see Table 1). To accommodate the broadest range of follow-up schedules, we focus on evaluating mediation using longitudinal data from two time points: (1) baseline and (2) the first post-baseline follow-up for which both mediation and outcome data were collected in each study.

Figure 1 depicts a basic two-wave longitudinal mediation model (MacKinnon, 2008; Valente & MacKinnon, 2017) that controls for baseline levels of both the mediator and the study outcome. This is an extension of the classic cross-sectional mediation model outlined by Baron and Kenny (1986) that evaluates if (a) the intervention (vs. control) is prospectively associated with post-baseline improvements in the mediator, (b) post-baseline improvement in the mediator is associated with post-baseline improvements in the study outcome, and (c) the intervention (vs. control) is associated with the study outcome after controlling for the mediator (i.e., the direct effect). This mediation model can be easily extended to include additional treatment contrasts and covariates as well as to accommodate clustered data across multiple studies, within an SEM framework. Next, we describe the application of the basic two-wave longitudinal mediation model outlined in Fig. 1 to the Project INTEGRATE

data. The meta-analytic mediation model consists of (1) an “overall model” that combines IPD across all studies and (2) “study-specific sub-models” that characterize potential differences between individual studies and inform the interpretation of the overall meta-analytic results.

Overall Mediation Model

First, we detail the overall meta-analytic mediation model of the combined sample of all participants across all included studies. Let $POST_PBS_{is}$ be the post-baseline PBS score of participant i in study s . Equation (1) is the first equation in the mediation model, which models the average, prospective association between each intervention group (vs. control) and post-baseline levels of the mediator variable, controlling for baseline levels of the mediator variable, PBS, and the study outcome variable, alcohol-related problems:

$$\begin{aligned} POST_PBS_{is} = & b_{0(A)} + b_{1(A)}TX_MIPF_{is} + b_{2(A)}TX_PF_{is} \\ & + b_{3(A)}TX_GMI_{is} + b_{4(A)}BL_ALCPROB_{is} \\ & + b_{5(A)}BL_PBS_{is} + b_{6(A)}MALE_{is} + b_{7(A)}FIRSTYR_{is} \\ & + b_{8(A)}NONWHITE_{is} + e_{is(A)}, \end{aligned} \quad (1)$$

where (A) identifies regression coefficients from the first of the two mediation model equations and $e_{is(A)}$ is a participant-specific residual error term. TX_MIPF_{is} , TX_PF_{is} , and TX_GMI_{is} are dummy-coded variables that indicate random allocation to MI + PF, PF, or GMI, respectively (each coded 1), compared to controls (all coded 0). The regression coefficients $b_{1(A)}$, $b_{2(A)}$, and $b_{3(A)}$ quantify the covariate-adjusted average difference between participants who received (1) MI + PF, (2) stand-alone PF, or (3) GMI, respectively, compared to control participants. The covariate BL_PBS_{is} adjusts for initial levels of the PBS mediator, and the covariate $BL_ALCPROB_{is}$ adjusts for initial levels of alcohol-related problems.

Let $POST_ALCPROB_{is}$ be the post-baseline level of the study outcome variable, alcohol-related problems, of participant i in study s . Equation (2) is the second equation in the mediation model, which models the association between post-baseline levels of the mediator, PBS, and post-baseline levels of the study outcome, alcohol-related problems, adjusting for baseline levels of the mediator and study outcome variables:

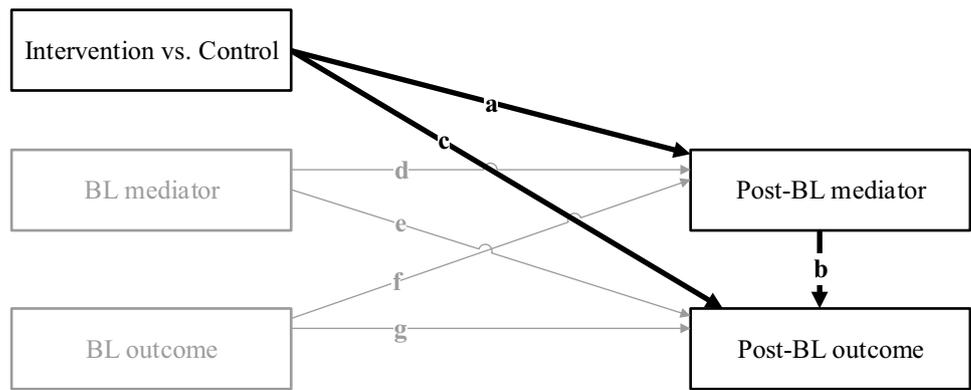
$$\begin{aligned} POST_ALCPROB_{is} = & b_{0(B)} + b_{1(B)}TX_MIPF_{is} + b_{2(B)}TX_PF_{is} \\ & + b_{3(B)}TX_GMI_{is} + b_{4(B)}BL_ALCPROB_{is} \\ & + b_{5(B)}BL_PBS_{is} + b_{6(B)}POST_PBS_{is} + b_{7(B)}MALE_{is} \\ & + b_{8(B)}FIRSTYR_{is} + b_{9(B)}NONWHITE_{is} + e_{is(B)}, \end{aligned} \quad (2)$$

where (B) identifies regression coefficients associated with the second mediation model equation and $e_{is(B)}$ is a participant-specific residual error term. The regression coefficients $b_{1(B)}$,

Table 2 Participant characteristics, proportions of missing data, and sample sizes by study

	Study										All studies (N = 3691)	
	2 (n = 144)	8a (n = 827)	8b (n = 1129)	8c (n = 244)	9 (n = 314)	12 (n = 157)	16 (n = 183)	18 (n = 140)	21 (n = 193)	22 (n = 360)		
Sex												
- Female	44 (30.6%)	576 (69.6%)	688 (60.9%)	147 (60.2%)	200 (63.7%)	83 (52.9%)	183 (100%)	101 (72.1%)	125 (64.8%)	207 (57.5%)	2354 (63.8%)	
- Male	100 (69.4%)	251 (30.4%)	441 (39.1%)	97 (39.8%)	114 (36.3%)	74 (47.1%)	-	38 (27.1%)	68 (35.2%)	153 (42.5%)	1336 (36.2%)	
- Missing	-	-	-	-	-	-	-	1 (0.7%)	-	-	1 (0.03%)	
Race												
- White	101 (70.1%)	713 (86.2%)	752 (66.6%)	210 (86.1%)	230 (73.2%)	143 (91.1%)	118 (64.5%)	127 (90.7%)	166 (86.0%)	330 (91.7%)	2890 (78.3%)	
- Non-White	40 (27.8%)	103 (12.5%)	369 (32.7%)	31 (12.7%)	84 (26.8%)	11 (7.0%)	63 (34.4%)	13 (9.3%)	27 (14.0%)	30 (8.3%)	771 (20.9%)	
- Missing	3 (2.1%)	11 (1.3%)	8 (0.7%)	8 (1.2%)	-	3 (1.9%)	2 (1.1%)	-	-	-	30 (0.8%)	
First-year student												
- No	59 (41.0%)	435 (52.6%)	592 (52.4%)	152 (62.3%)	-	152 (96.8%)	-	98 (70.0%)	115 (59.6%)	-	1603 (43.4%)	
- Yes	85 (59.0%)	386 (46.7%)	532 (47.1%)	91 (37.3%)	314 (100%)	5 (3.2%)	183 (100%)	41 (29.3%)	78 (40.4%)	360 (100%)	2075 (56.2%)	
- Missing	-	6 (0.7%)	5 (0.4%)	1 (0.4%)	-	-	-	1 (0.7%)	-	-	13 (0.4%)	
Alcohol-related problems												
- Baseline												
- Mean (SD)	-0.5 (0.7)	0.1 (0.9)	0.2 (1.0)	0.1 (0.9)	0.9 (0.7)	0.4 (0.6)	-0.1 (0.9)	0.1 (0.9)	0.9 (0.8)	0.0 (0.9)	0.2 (0.9)	
- Range	[-1.7, 1.2]	[-1.5, 3.5]	[-1.7, 3.5]	[-1.5, 2.6]	[-1.2, 2.9]	[-1.2, 1.8]	[-1.7, 1.9]	[-1.6, 2.1]	[-0.7, 2.9]	[-1.5, 2.0]	[-1.7, 3.5]	
- Missing	-	3 (0.4%)	3 (0.3%)	-	-	-	-	-	-	-	6 (0.2%)	
- Post-baseline												
- Mean (SD)	-0.7 (0.7)	0.2 (0.8)	0.3 (0.9)	0.2 (0.9)	0.8 (0.8)	0.3 (0.6)	0.1 (1.0)	0.0 (1.0)	0.6 (0.8)	0.3 (0.9)	0.3 (0.9)	
- Range	[-1.7, 1.3]	[-1.4, 3.3]	[-1.7, 3.3]	[-1.5, 3.2]	[-1.3, 3.6]	[-1.2, 1.8]	[-1.7, 2.6]	[-1.6, 2.9]	[-0.9, 3.3]	[-1.5, 2.4]	[-1.7, 3.6]	
- Missing	-	-	-	-	-	-	-	2 (1.4%)	2 (1.0%)	-	4 (0.1%)	
Protective behavioral strategies												
- Baseline												
- Mean (SD)	0.9 (0.8)	0.5 (0.8)	0.5 (0.9)	0.4 (0.9)	0.4 (0.8)	-0.4 (0.9)	0.8 (1.2)	0.5 (0.8)	0.3 (0.9)	0.4 (0.8)	0.5 (0.9)	
- Range	[-1.0, 3.1]	[-2.3, 2.9]	[-2.2, 3.4]	[-2.4, 3.2]	[-1.6, 2.6]	[-2.3, 2.7]	[-2.8, 3.2]	[-2.0, 3.4]	[-2.2, 3.4]	[-2.4, 2.6]	[-2.8, 3.4]	
- Missing	18 (12.5%)	206 (24.9%)	258 (22.9%)	40 (16.4%)	-	-	66 (36.1%)	9 (6.4%)	-	72 (20.0%)	669 (18.1%)	
- Post-baseline												
- Mean (SD)	1.0 (0.9)	0.6 (0.8)	0.6 (0.9)	0.6 (0.9)	0.4 (0.8)	-0.3 (0.9)	0.7 (1.1)	0.6 (1.2)	0.4 (0.9)	0.3 (0.9)	0.5 (0.9)	
- Range	[-1.7, 3.4]	[-1.6, 3.2]	[-2.6, 3.5]	[-2.4, 2.6]	[-1.7, 2.9]	[-2.3, 1.6]	[-2.4, 2.8]	[-2.8, 3.4]	[-2.0, 3.4]	[-2.6, 2.6]	[-2.8, 3.5]	
- Missing	0 (0.0%)	12 (1.5%)	25 (2.2%)	1 (0.4%)	6 (1.9%)	-	3 (1.6%)	2 (1.4%)	2 (1.0%)	-	51 (1.4%)	

Fig. 1 A two-wave longitudinal intervention mediation model. The key mediation-related paths are emphasized in black and the covariate paths are in gray. BL, baseline; post-BL, post-baseline



$b_{2(B)}$, and $b_{3(B)}$ provide the average *direct effect* of each intervention (vs. control) across studies on post-baseline levels of the study outcome variable. Both Eqs. (1) and (2) include the demographic covariates $MALE_{is}$, $FIRSTYR_{is}$, and $NONWHITE_{is}$, which adjust for sex (1 = men vs. 0 = women), first-year student status (1 = first-year vs. 0 = non-first-year), and race (1 = non-White vs. 0 = White), respectively.

Study-Specific Mediation Sub-models

Next, we describe the study-specific mediation sub-models, which inform the interpretation of the overall mediation model by characterizing variation in the results across studies. The mediation analysis is repeated separately and sequentially for each study by using sub-models of Eqs. (1) and (2) to include the estimable terms (i.e., evaluated intervention groups and demographic covariates with variability). For example, coefficients $b_{7(A)}$ and $b_{8(B)}$ are not estimable and hence excluded in the study-specific sub-models for studies 9, 16, and 22 because they recruited only first-year students. As an illustration, Eqs. (3) and (4) are the study-specific sub-models for study 22, which evaluated MI + PF vs. control:

$$POST_PBS_i = b_{0(A)} + b_{1(A)}TX_MIPF_i + b_{4(A)}BL_ALCPROB_i + b_{5(A)}BL_PBS_i + b_{6(A)}MALE_i + b_{8(A)}NONWHITE_i + e_{i(A)}, \tag{3}$$

$$POST_ALCPROB_i = b_{0(B)} + b_{1(B)}TX_MIPF_i + b_{4(B)}BL_ALCPROB_i + b_{5(B)}BL_PBS_i + b_{6(B)}POST_PBS_i + b_{7(B)}MALE_i + b_{9(B)}NONWHITE_i + e_{i(B)}, \tag{4}$$

where (A) and (B) identify regression coefficients from the reduced first and second mediation model equations, respectively, i identifies the participant, and $e_{i(A)}$ and $e_{i(B)}$ are participant-specific residual error terms. For consistency, the subscripts in Eqs. (3) and (4) correspond with the same variables as those shown in the overall model Eqs. (1) and (2). As seen in Table 1, intervention groups not evaluated in a study become study-level missing data in the context of IPD

meta-analysis. The parameters associated with missing treatment contrasts are excluded in the study-specific sub-model. Thus, $b_{2(A)}$, $b_{2(B)}$, $b_{3(A)}$, and $b_{3(B)}$ are excluded from Eqs. (3) and (4) since PF and GMI were not evaluated in study 22 by study design (i.e., $TX_MIPF_i = 0$ and $TX_GMI_i = 0$ for all participants i in study s), and $b_{7(A)}$, $b_{8(B)}$ are excluded by study design since all participants in study 22 were first-year students. It is important to note that the interpretation of each parameter estimate depends on the other parameters included in the model (see Jiao et al., 2020). However, if we assume that Eqs. (1) and (2) represent the true model for all studies, it is reasonable to assume the omitted coefficients in the sub-models are missing at random. In addition, since baseline PBS and alcohol-related problems are adjusted for in all sub-models, any interpretational bias associated with missing demographic covariates would be minimal.

Accounting for Clustered Design Using SEM for Complex Survey Data

A key data feature of IPD combined from multiple studies is the nesting of individual participants within studies, which must be considered for accurate statistical inference (see also Mun et al., 2015, p. 36–38). To account for the nested data structure of IPD from multiple studies in a one-stage integrative analysis, parameter estimates and corresponding standard errors can be adjusted for clustering by utilizing either (1) a model-based approach using multilevel modeling that incorporates cluster-specific parameters (e.g., Huh et al., 2015, 2019) or (2) a design-based approach in which clustering is accommodated via complex survey analysis with weights applied to participants in a single-level analysis (e.g., Clarke et al., 2013, 2016; Li et al., 2020; Ray et al., 2014).

The advantage of design-based adjustment for clustering is that it can be implemented easily in an SEM framework and produces estimates that are comparable to multilevel modeling (Wu & Kwok, 2012), but with a lower

computational burden. The computational efficiency of cluster-adjusted SEM makes it especially useful when combined with bootstrapping, the commonly accepted method for evaluating the statistical significance of the mediated (i.e., indirect) effect (see “Bootstrap Resampling with Multiple Imputation” later).

To evaluate the meta-analytic mediation model outlined in Eqs. (1)–(4), while accounting for the nested design of the data, we utilized SEM for complex survey data by first using the R package lavaan (Rosseel, 2012) to estimate an SEM that combines data across all studies in a single-level analysis followed by lavaan.survey (Oberski, 2014), which provides a design-based adjustment to account for clustering by study. SEM for complex survey data is analogous to the generalized estimating equation (Zeger et al., 1988) approach to analyzing multilevel data, which is also a design-based approach to accommodate clustered data. To account for widely varying sample sizes across studies, we weighted the data using the inverse of the square root of each study’s sample size as explained in Mun et al. (2015) and used in research applications (Clarke et al., 2013, 2016; Ray et al., 2014).

With respect to interpretation, the regression coefficients (i.e., fixed effects) produced by SEM for complex survey data are marginal estimates, which represent the average effects across all individuals. In contrast, regression coefficients estimated using a model-based approach are cluster-specific estimates that are conditional on specific values of the random effects (e.g., the deviation of a specific individual from the group average). When the outcome is modeled as normally distributed, regression coefficients produced by multilevel models (i.e., mixed-effects models) can be interpreted like marginal estimates, although this does not hold for extensions of multilevel modeling that use a non-identity link function, such as logistic or Poisson models (Atkins et al., 2013). Thus, the inference for a model that accounts for clustering using a design-based approach is functionally equivalent to multilevel modeling in the present application.

Calculating the Indirect and Total Effect of Intervention

To calculate the *indirect effect* of each intervention type on the post-baseline study outcome via changes in the mediator, we calculate the product of the regression coefficients corresponding to (1) the association between intervention type and post-baseline PBS (i.e., $b_{1(A)}$, $b_{2(A)}$, and $b_{3(A)}$) and (2) the association between post-baseline PBS and changes in alcohol-related problems, $b_{6(B)}$. Equations (5)–(7) summarize the formulas used to calculate the indirect effects of MI + PF, stand-alone PF, and GMI vs. control, respectively,

for the *overall* (Eqs. 1 and 2) and *study-specific* (Eqs. 3 and 4) models:

$$\text{MI+PF} : b_{1(A)} \times b_{6(B)}, \quad (5)$$

$$\text{Stand-alone PF} : b_{2(A)} \times b_{6(B)}, \text{ and} \quad (6)$$

$$\text{GMI} : b_{3(A)} \times b_{6(B)}. \quad (7)$$

To calculate the *total effect* of each intervention type on post-baseline alcohol-related problems, we sum (a) the direct effect of each intervention type on alcohol-related problems (i.e., $b_{1(B)}$, $b_{2(B)}$, and $b_{3(B)}$) from Eq. (2) and (b) the corresponding indirect effect of each intervention type calculated in Eqs. (5), (6), or (7). Equations (8)–(10) summarize the formulas used to calculate the total effects of MI + PF, stand-alone PF, and GMI vs. control, respectively, for the *overall* (Eqs. 1 and 2) and *study-specific* (Eqs. 3 and 4) models:

$$\text{MI+PF} : b_{1(B)} + (b_{1(A)} \times b_{6(B)}), \quad (8)$$

$$\text{Stand-alone PF} : b_{2(B)} + (b_{2(A)} \times b_{6(B)}), \text{ and} \quad (9)$$

$$\text{GMI} : b_{3(B)} + (b_{3(A)} \times b_{6(B)}). \quad (10)$$

Bootstrap Resampling with Multiple Imputation

To evaluate the magnitude and statistical significance of the estimates from the mediation model, including regression coefficients, indirect effects, total effects, and R^2 values, we used bootstrap resampling (Efron & Tibshirani, 1993) in which the mediation analyses are replicated across 5000 bootstrapped data sets to calculate the mean point estimate and 95% confidence interval for each parameter. Bootstrap estimation involves random sampling of observations with replacement from the original data set such that the sample is treated as if it were the population. The effect of sampling with replacement is that an observation may be represented more than once, whereas some observations may be left out in any given bootstrap sample. As a result, the bootstrap sample is equal in size to the original but is not identical.

Because missing data present in the original data set will also be reflected in the bootstrap data set, an additional consideration is needed to handle missing data when bootstrapping. In the context of an IPD meta-analysis, there can be two sources of missing data: (1) study-level missing data due to a variable not being assessed or without variation (see Jiao et al., 2020; Kim et al., 2014) and (2) participant-level

missing data due to nonresponse. In the context of the Project INTEGRATE data, study-level missing data occurred because only one study evaluated all three intervention groups, the rest evaluated a subset of intervention groups (i.e., one or two), and also because some studies exclusively targeted first-year students or women. These are not missing variables within the original studies; however, in the context of meta-analysis, they are missing or inestimable covariates at the study level. As described previously, we excluded the corresponding treatment contrast or demographic covariate from the corresponding study-specific mediation sub-model. Therefore, study-level missing variables were not imputed.

As seen in Table 2, there were also participant-level missing variables. Thus, to minimize bias in the results of the mediation analysis due to missing mediator, outcome, and/or covariate data, bootstrapping was combined with multiple imputation. Multiple imputation is a widely used method for accommodating missing data. Furthermore, simulation research supports combining multiple imputation with bootstrapping (Little & Rubin, 2002; Schomaker & Heumann, 2018). There are several ways to combine multiple imputation and bootstrapping, each with pros and cons (Brand et al., 2019). In the present study, we chose to bootstrap first, followed by multiple imputation, which is more computationally intensive but produces confidence intervals that more accurately reflect uncertainty due to missing data (Bartlett & Hughes, 2020).

First, a stratified bootstrap was performed in which participants, including those with missing data, were randomly sampled with replacement separately by study and intervention group then combined into a single bootstrapped data set of equal size to the original data set. The stratification by study and intervention group accounted for the clustered design (i.e., participants nested within studies and groups) and maintained consistent sample sizes in subsequent analyses, within and across studies, as well as across all intervention groups. A total of 5000 bootstrap-resampled data sets were generated. Second, for each of the bootstrap-resampled data sets, a set of ten imputed data sets were generated via multivariate normal imputation with the R package Amelia (Honaker et al., 2011). According to simulation findings by Bartlett and Hughes (2020), ten imputations per bootstrap replicate provide approximately accurate confidence intervals when multiple imputation is nested within bootstrapping.

The mediation analysis was repeated for each multiply imputed data set, and the results were combined across ten imputed data sets. This yielded a set of 5000 estimates for each parameter in the mediation model, one for each bootstrap replicate. The collection of bootstrap estimates approximates the sampling distribution for each parameter and accommodates non-normally distributed estimates, such as the indirect and total effects. The point estimate for each

parameter was calculated as the mean across the 5000 bootstrap replications. Bias-corrected and accelerated 95% confidence intervals were calculated to assess the indirect and total effects, as recommended by MacKinnon et al. (2004).

Analysis of the Motivating Data and the Summary of Findings

Annotated computer code in R for fitting the model, along with example data, can be accessed in the online repository (<https://doi.org/10.17632/t2yk5kt3bw.1>; Huh et al., 2021).

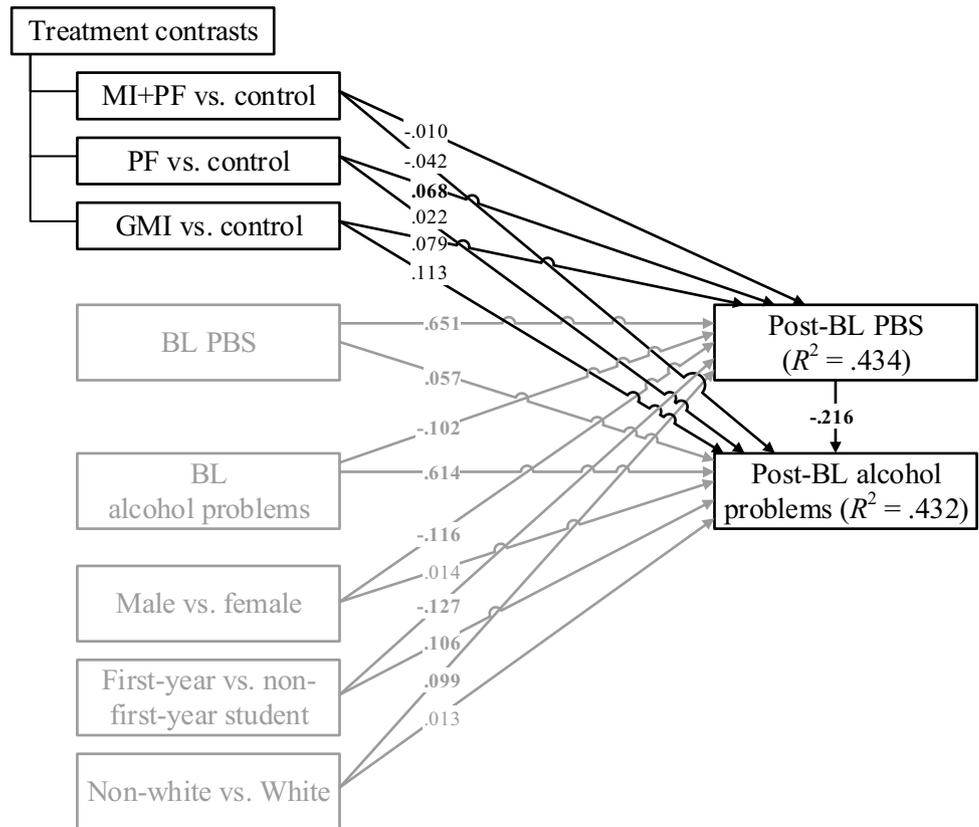
Figure 2 is a path diagram that summarizes the estimated associations from the overall mediation model of the combined sample. The path coefficients are standardized with respect to the outcome, which can be interpreted as the effect that a unit difference in each predictor has on the corresponding outcome variable, holding all other covariates constant. For treatment contrasts and other indicator variables, the path coefficients correspond with the difference between groups (e.g., MI + PF vs. control) in *SDs* of the outcome. For continuous predictors (i.e., alcohol-related problems, PBS), the standardized coefficient can be interpreted as the change in *SDs* of the outcome for a unit difference in the predictor. The overall mediation model explained 43% of the variance in both post-baseline PBS and post-baseline alcohol-related problems.

The paths of interest are (1) the prospective association between each intervention and post-baseline levels of the mediator (PBS) and (2) the association of the mediator and the outcome at post-baseline. Of the three interventions, only stand-alone PF had a statistically significant association with the mediator, with a .07 *SD* increase (95% CI = [.01, .12]) in post-baseline PBS as compared to control. A one-*SD* increase in post-baseline PBS, in turn, was associated with a .22 *SD* reduction (95% CI = [−.26, −.17]) in post-baseline alcohol-related problems.

Figure 3 is a forest plot that summarizes the key mediation-related results (i.e., indirect and total effects) from (a) the ten study-specific sub-models (top portion) and (b) the overall model (bottom portion, highlighted in gray) of the combined sample. A negative coefficient can be interpreted as a prospective improvement (i.e., reduction) in alcohol-related problems at post-baseline. Stand-alone PF, compared with control, was associated with a statistically significant, albeit small, reduction in alcohol-related problems via increased use of PBS ($\beta = -.01$, 95% CI = [−.03, −.002]). Neither MI + PF nor GMI was associated with statistically significant reductions in alcohol-related problems, compared with control, through improvements in PBS.

An additional sensitivity analysis was conducted to evaluate the consistency of the findings when the mediation analysis was repeated by leaving out one study at a time,

Fig. 2 Overall mediation model evaluating change in protective behavioral strategies as a pathway by which brief motivational intervention improves alcohol-related problems for college students who drink. The key mediation-related paths are emphasized in black and the covariate paths are in gray. All path coefficients are standardized betas (with respect to the outcomes only), and results highlighted in bold are statistically significant ($p < .05$). MI + PF, individually delivered motivational interviewing with personalized feedback; PF, stand-alone personalized feedback; GMI, group motivational interviewing; BL, baseline; post-BL, post-baseline; PBS, protective behavioral strategies



sequentially (see the [Supplemental Material](#) for a summary). The indirect and total effects of each intervention approach were consistent across the sensitivity models, suggesting

that the results were robust and not driven by any single influential study.

Study	Intervention	Indirect Effect		Total Effect	
		Std Beta	[95% CI]	Std Beta	[95% CI]
2	PF	-.034	[-.143, .052]	.127	[-.208, .447]
8a	PF	-.012	[-.042, .015]	-.012	[-.042, .015]
8b	PF	-.011	[-.028, .005]	-.002	[-.096, .088]
8c	PF	.035	[.002, .101]	.013	[-.181, .208]
9	MI+PF	-.016	[-.067, .011]	.078	[-.162, .315]
	PF	.011	[-.021, .050]	.109	[-.096, .327]
	GMI	.028	[-.003, .077]	.147	[-.080, .379]
12	MI+PF	-.006	[-.059, .032]	-.416	[-.629, -.188]
16	GMI	-.021	[-.088, .003]	.179	[-.037, .384]
18	PF	-.057	[-.207, .083]	.251	[-.060, .566]
21	MI+PF	-.067	[-.165, .003]	-.087	[-.345, .188]
	PF	-.017	[-.125, .069]	.066	[-.207, .341]
22	MI+PF	-.018	[-.067, .009]	-.031	[-.202, .137]
Overall Model	MI+PF	.002	[-.019, .019]	-.040	[-.126, .042]
	PF	-.015	[-.027, -.002]	.007	[-.051, .064]
	GMI	-.017	[-.049, .015]	.096	[-.037, .228]

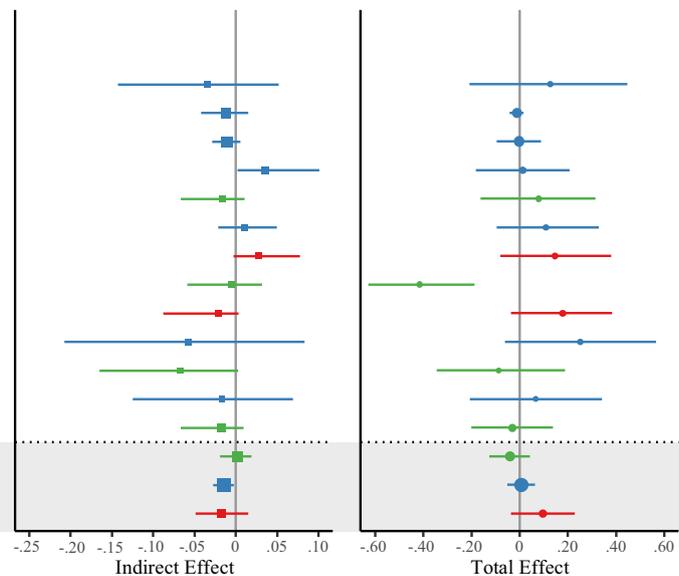


Fig. 3 Forest plot of the indirect and total effect of brief motivational interventions on alcohol-related problems via protective behavioral strategies. The study-specific sub-model results (top portion) illustrate the heterogeneity in the indirect and total effects of intervention across the ten studies. The overall model results (bottom portion, highlighted in gray) for the combined sample summarize the *indirect*

and *total effects* of each intervention type on post-baseline alcohol-related problems. Std Beta, standardized beta, with respect to the outcome only; CI, confidence interval; MI + PF, individually delivered motivational interviewing with personalized feedback (green-colored estimates); PF, stand-alone personalized feedback (blue-colored estimates); GMI, group motivational interviewing (red-colored estimates)

Discussion

The literature evaluating mechanisms of intervention effect has relied almost exclusively on single-study intervention trials, which are frequently underpowered to evaluate mediation hypotheses (Fritz et al., 2015). This methodological illustration details a meta-analytic mediation analysis approach that leverages IPD across multiple studies to evaluate mechanisms of change longitudinally. Specifically, the approach evaluates whether the prospective change in a mediator following intervention is accompanied by a change in the outcome. Moreover, the approach can accommodate missing data commonly encountered in clinical trial data, making it a practical option for meta-analytic mediation analysis.

The illustrated SEM approach combines well-established quantitative methodologies, including SEM with design-based adjustment for clustering, bootstrap estimation of mediated effects, and multiple imputation, to test mediation with accuracy and precision. We describe how to calculate the magnitude of a mediated effect within and across studies and assess its statistical significance in a way that (a) accounts for the clustering of participants within the study, (b) uses all available data, and (c) produces point estimates and confidence intervals for the indirect and total effects of an intervention that account for the non-normal distribution that arises from a product of coefficients.

At a substantive level, it is of interest that greater use of PBS mediated the effect of stand-alone PF intervention on alcohol-related problems. Specifically, participants receiving stand-alone PF had greater improvement in PBS utilization compared with participants randomized to the control comparison. Greater PBS utilization, in turn, was associated with concurrent reductions in alcohol-related problems. Although statistically significant, it is important to note that the mediated effect of PF via a change in PBS was quite small, equivalent to a .01 *SD* difference in the reduction in alcohol-related problems. The small mediated effect may be because brief motivational interventions, including PF, do not increase the use of PBS substantially. However, the results from this study may suggest that stand-alone PF focusing on a few salient points, such as PBS, may be more likely to induce behavior change than formats that use multiple modalities (Ray et al., 2014).

Although the effect of brief motivational interventions on alcohol-related problems via a change in PBS appeared to be quite small in the present study, our findings are consistent with the evidence of some PBS-based interventions failing to improve outcomes (Martens et al., 2013). In addition, college students utilize PBS for different reasons, with some students engaging in PBS to get intoxicated faster while trying to prevent the most extreme harm. Therefore, the increased use of PBS can increase alcohol-related problems for some students unmotivated to change their drinking,

while low-risk drinkers may use them to effectively limit harm from drinking (Li et al., 2020). The average effect that we focused on in the current study, although important, needs to be examined further for heterogeneous mediational paths, accounting for students' different motivations for drinking and PBS use.

It is important to note that most of the studies evaluated only one or two intervention groups and not all three interventions. The unbalanced nature of the intervention groups across studies is a typical challenge in a meta-analysis across heterogeneous studies, including IPD data syntheses (Brincks et al., 2018; Huh et al., 2019), and can complicate the interpretation of findings. However, the motivating data featured a large, pooled sample of college students from brief motivational intervention studies, which permitted more robust mediation estimates for all the intervention types (i.e., MI + PF, stand-alone PF, and GMI) than would be possible in individual trials. Furthermore, we previously developed commensurate measures across trials for key constructs and carefully controlled for baseline levels of both the mediator and outcome variables, which bolsters confidence in the findings.

An important advantage of meta-analytic mediation analysis using IPD compared to traditional meta-analysis is the ability to evaluate the prospective association between baseline participant characteristics and change in PBS, which yielded additional insights. As seen in Fig. 2, we found that men (vs. women), first-year students (vs. non-first-year students), White students (vs. non-White students), and those with more severe alcohol-related problems at baseline showed less improvement in PBS use at a follow-up. The ability to make inferences regarding participant-level change shows the benefit of this IPD-based approach for evaluating mechanisms of change in prevention research.

Limitations and Future Directions

It is important to consider the limitations of the present study. First, we could not evaluate if change in the mediator *preceded* change in the study outcome, which would require data from at least three time points. Second, the approach relies on assumptions about missing data that we believe to be reasonable, including that the absence of an intervention group in a study does not bias the overall findings. However, further investigation via simulation study may be needed to identify potential areas of improvement. Third, this methodological illustration focuses on evaluating a single mediator; however, the approach we detailed can be extended to models with multiple mediators. Fourth, a minor drawback to our approach is that combining multiple imputation with bootstrapping is computationally intensive; however, the estimation times (e.g., 10–20 min per model) encountered in the present study are feasible for applied research. Finally, our motivating example focused on a relatively normally distributed mediator variable

and outcome of interest. Future research might examine extensions of this approach within a generalized SEM framework to binary, count, or other outcome distributions.

Conclusions

The SEM approach detailed in this methodological illustration is a flexible approach for conducting a mediation analysis that leverages the most granular information from multiple studies and overcomes key challenges that arise when combining clinical trial data. The annotated R code and data provide additional guidance for researchers who wish to apply the method in their own research, and we hope it will motivate further development in meta-analytic mediation methodology and its applications in prevention science.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11221-021-01318-4>.

Acknowledgements We would like to thank the following contributors to Project INTEGRATE in alphabetical order: John S. Baer, Department of Psychology, The University of Washington, and Veterans' Affairs Puget Sound Health Care System; Nancy P. Barnett, Center for Alcohol and Addiction Studies, Brown University; M. Dolores Cimini, University Counseling Center, The University at Albany, State University of New York; William R. Corbin, Department of Psychology, Arizona State University; Kim Fromme, Department of Psychology, The University of Texas at Austin; Joseph W. LaBrie, Department of Psychology, Loyola Marymount University; Mary E. Larimer, Department of Psychiatry and Behavioral Sciences, The University of Washington; Matthew P. Martens, Department of Educational, School, and Counseling Psychology, The University of Missouri; James G. Murphy, Department of Psychology, The University of Memphis; Scott T. Walters, Department of Health Behavior and Health Systems, The University of North Texas Health Science Center; Helene R. White, Center of Alcohol and Substance Use Studies, Rutgers, The State University of New Jersey; and the late Mark D. Wood, Department of Psychology, The University of Rhode Island. We would like to thank Minge Xie, Department of Statistics, Rutgers University, and Jae-kwang Kim, Department of Statistics, Iowa State University, for their suggestions on statistical issues. We also thank Nিকেisha Clarke, Yang Jiao, Su-Young Kim, and Anne E. Ray for their earlier work on coding and harmonizing interventions and outcomes, and Jimmy de la Torre and Yan Huo for their work on measurement. Finally, we thank Helene R. White for her valuable conceptual and methodological contributions in the early years of Project INTEGRATE.

Funding The project described was supported by the National Institute on Alcohol Abuse and Alcoholism (NIAAA) grants R01 AA019511 and K02 AA028630. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIAAA or the National Institutes of Health.

Declarations

Research Involving Human Participants This project was approved by the North Texas Regional Institutional Review Board (IRB). The original trials that contributed to Project INTEGRATE were IRB approved in each of the respective institutions. All ethical standards

for conducting research with human participants were followed in the current project as well as in the implementation of the original trials, including the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards.

Informed Consent Informed consent was obtained from all participants included in the original studies contributing to this meta-analysis.

Conflict of Interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

Studies included in the meta-analysis are marked in the References with an asterisk.

- American College Health Association. (2001). *National College Health Assessment ACHA-NCHA reliability and validity analyses*. Baltimore, MD: American College Health Association.
- Atkins, D. C., Baldwin, S. A., Zheng, C., Gallop, R. J., & Neighbors, C. (2013). A tutorial on count regression and zero-altered count models for longitudinal substance use data. *Psychology of Addictive Behaviors*, 27, 166–177. <https://doi.org/10.1037/a0029508>
- Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173–1182. <https://doi.org/10.1037/0022-3514.51.6.1173>
- Bartlett, J. W., & Hughes, R. A. (2020). Bootstrap inference for multiple imputation under uncongentiality and misspecification. *Statistical Methods in Medical Research*, 29, 3457–3491. <https://doi.org/10.1177/0962280220932189>
- Brand, J., van Buuren, S., le Cessie, S., & van den Hout, W. (2019). Combining multiple imputation and bootstrap in the analysis of cost-effectiveness trial data. *Statistics in Medicine*, 38, 210–220. <https://doi.org/10.1002/sim.7956>
- Brincks, A., Montag, S., Howe, G. W., Huang, S., Siddique, J., Ahn, S., et al. (2018). Addressing methodologic challenges and minimizing threats to validity in synthesizing findings from individual-level data across longitudinal randomized trials. *Prevention Science*, 19, 60–73. <https://doi.org/10.1007/s11221-017-0769-1>
- Cheung, M. W. -L. (2014). Fixed- and random-effects meta-analytic structural equation modeling: Examples and analyses in R. *Behavior Research Methods*, 46, 29–40. <https://doi.org/10.3758/s13428-013-0361-y>

- Cheung, M. W. -L. (2015). *Meta-analysis: A structural equation modeling approach*. Wiley.
- Cheung, M. W. -L., & Chan, W. (2005). Meta-analytic structural equation modeling: A two-stage approach. *Psychological Methods*, 10, 40–64. <https://doi.org/10.1037/1082-989X.10.1.40>
- Cheung, M. W. -L., & Chan, W. (2009). A two-stage approach to synthesizing covariance matrices in meta-analytic structural equation modeling. *Structural Equation Modeling*, 16, 28–53. <https://doi.org/10.1080/10705510802561295>
- Clarke, N., Kim, S. -Y., Ray, A. E., White, H. R., Jiao, Y., & Mun, E. -Y. (2016). The association between protective behavioral strategies and alcohol-related problems: An examination of race and gender differences among college drinkers. *Journal of Ethnicity in Substance Abuse*, 15, 25–45. <https://doi.org/10.1080/15332640.2014.1002877>
- Clarke, N., Kim, S. -Y., White, H. R., Jiao, Y., & Mun, E. -Y. (2013). Associations between alcohol use and alcohol-related negative consequences among Black and White college men and women. *Journal of Studies on Alcohol and Drugs*, 74, 521.
- Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. Chapman & Hall.
- Fritz, M. S., Cox, M. G., & MacKinnon, D. P. (2015). Increasing statistical power in mediation models without increasing sample size. *Evaluation & the Health Professions*, 38, 343–366. <https://doi.org/10.1177/0163278713514250>
- Honaker, J., King, G., & Blackwell, M. (2011). Amelia II: A program for missing data. *Journal of Statistical Software*, 45(1), 1–47. <https://doi.org/10.18637/jss.v045.i07>
- Huh, D., Li, X., Zhou, Z., Walters, S. T., Baldwin, S. A., Tan, Z., et al. (2021). Data and code for: Huh et al. (2021). A structural equation modeling approach to meta-analytic mediation analysis using individual participant data. *Mendeley Data*, v1. <https://doi.org/10.17632/t2yk5kt3bw.1>
- Huh, D., Mun, E. -Y., Larimer, M. E., White, H. R., Ray, A. E., Rhew, I. C., et al. (2015). Brief motivational interventions for college student drinking may not be as powerful as we think: An individual participant-level data meta-analysis. *Alcoholism: Clinical and Experimental Research*, 39(5), 919–931. <https://doi.org/10.1111/acer.12714>
- Huh, D., Mun, E. -Y., Walters, S. T., Zhou, Z., & Atkins, D. C. (2019). A tutorial on individual participant data meta-analysis using Bayesian multilevel modeling to estimate alcohol intervention effects across heterogeneous studies. *Addictive Behaviors*, 94, 162–170. <https://doi.org/10.1016/j.addbeh.2019.01.032>
- Huo, Y., de la Torre, J., Mun, E. -Y., Kim, S. -Y., Ray, A. E., Jiao, Y., & White, H. R. (2015). A hierarchical multi-unidimensional IRT approach for analyzing sparse, multi-group data for integrative data analysis. *Psychometrika*, 80, 834–855. <https://doi.org/10.1007/s11336-014-9420-2>
- Hurlbut, S. C., & Sher, K. J. (1992). Assessing alcohol problems in college students. *Journal of American College Health*, 41, 49–58. <https://doi.org/10.1080/07448481.1992.10392818>
- Hussong, A. M., Curran, P. J., & Bauer, D. J. (2013). Integrative data analysis in clinical psychology research. *Annual Review of Clinical Psychology*, 9, 61–89. <https://doi.org/10.1146/annurev-clinpsy-050212-185522>
- Jiao, Y., Mun, E. -Y., Trikalinos, T. A., & Xie, M. (2020). A CD-based mapping method for combining multiple related parameters from heterogeneous intervention trials. *Statistics and Its Interface*, 13, 533–549. <https://doi.org/10.4310/SII.2020.v13.n4.a10>
- Kahler, C. W., Strong, D. R., & Read, J. P. (2005). Toward efficient and comprehensive measurement of the alcohol problems continuum in college students: The Brief Young Adult Alcohol Consequences Questionnaire. *Alcoholism: Clinical and Experimental Research*, 29(7), 1180–1189. <https://doi.org/10.1097/01.ALC.0000171940.95813.A5>
- Kim, S. -Y., Mun, E. -Y., & Smith, S. (2014). Using mixture models with known class membership to address incomplete covariance structures in multiple-group growth models. *British Journal of Mathematical and Statistical Psychology*, 67, 94–116. <https://doi.org/10.1111/bmsp.12008>
- *LaBrie, J. W., Huchting, K. K., Lac, A., Tawalbeh, S., Thompson, A. D., & Larimer, M. E. (2009). Preventing risky drinking in first-year college women: Further validation of a female-specific motivational-enhancement group intervention. *Journal of Studies on Alcohol and Drugs, Supplement*, 16, 77–85. <https://doi.org/10.15288/jsads.2009.s16.77>
- *Larimer, M. E., Lee, C. M., Kilmer, J. R., Fabiano, P. M., Stark, C. B., Geisner, I. M., et al. (2007). Personalized mailed feedback for college drinking prevention: A randomized clinical trial. *Journal of Consulting and Clinical Psychology*, 75, 285–293. <https://doi.org/10.1037/0022-006X.75.2.285>
- *Lee, C. M., Kaysen, D. L., Neighbors, C., Kilmer, J. R., & Larimer, M. E. (2009). Feasibility, acceptability, and efficacy of brief interventions for college drinking: Comparison of group, individual, and web-based alcohol prevention formats. [Unpublished manuscript]. Department of Psychiatry and Behavioral Sciences, University of Washington.
- Li, X., Clarke, N., Kim, S. -Y., Ray, A. E., Walters, S. T., & Mun, E. -Y. (2020). Protective behavioral strategies are more helpful for avoiding alcohol-related problems for college drinkers who drink less. *Journal of American College Health*. Advance online publication. <https://doi.org/10.1080/07448481.2020.1807555>
- Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data* (2nd ed.). Wiley-Interscience.
- MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis*. Routledge.
- MacKinnon, D. P., Lockwood, C. M., & Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate Behavioral Research*, 39, 99. https://doi.org/10.1207/s15327906mbr3901_4
- Martens, M. P., Ferrier, A. G., Sheehy, M. J., Corbett, K., Anderson, D. A., & Simmons, A. (2005). Development of the protective behavioral strategies survey. *Journal of Studies on Alcohol*, 66(5), 698–705. <https://doi.org/10.15288/jsa.2005.66.698>
- *Martens, M. P., Kilmer, J. R., Beck, N. C., & Zamboanga, B. L. (2010). The efficacy of a targeted personalized drinking feedback intervention among intercollegiate athletes: A randomized controlled trial. *Psychology of Addictive Behaviors*, 24, 660–669. <https://doi.org/10.1037/a0020299>
- Martens, M. P., Smith, A. E., & Murphy, J. G. (2013). The efficacy of single-component brief motivational interventions among at-risk college drinkers. *Journal of Consulting and Clinical Psychology*, 81, 691–701. <https://doi.org/10.1037/a0032235>
- Mun, E. -Y., de la Torre, J., Atkins, D. C., White, H. R., Ray, A. E., Kim, S. -Y., et al. (2015). Project INTEGRATE: An integrative study of brief alcohol interventions for college students. *Psychology of Addictive Behaviors*, 29, 34–48. <https://doi.org/10.1037/adb0000047>
- Mun, E. -Y., Huo, Y., White, H. R., Suzuki, S., & de la Torre, J. (2019). Multivariate higher-order IRT model and MCMC algorithm for linking individual participant data from multiple studies. *Frontiers in Psychology*, 10, 1328. <https://doi.org/10.3389/fpsyg.2019.01328>
- Mun, E. -Y., Jiao, Y., & Xie, M. (2016). Integrative data analysis for research in developmental psychopathology. In D. Cicchetti (Ed.), *Developmental psychopathology: Theory and method* (3rd ed., Vol. 1, pp. 1042–1087). Hoboken, NJ: Wiley. <https://doi.org/10.1002/9781119125556>
- Mun, E. -Y., Li, X., Lineberry, S., Tan, Z., Huh, D., Walters, S. T., et al. (2021). Do brief alcohol interventions reduce driving after

- drinking among college students? A two-step meta-analysis of individual participant data. *Alcohol and Alcoholism*, agaa146. <https://doi.org/10.1093/alcalc/agaa146>
- Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement*, 16, 159–176. <https://doi.org/10.1177/014662169201600206>
- Oberski, D. (2014). lavaan.survey: An R package for complex survey analysis of structural equation models. *Journal of Statistical Software*, 57(1), 1–27. <https://doi.org/10.18637/jss.v057.i01>
- Ray, A. E., Kim, S. -Y., White, H. R., Larimer, M. E., Mun, E. -Y., Clarke, N., et al. (2014). When less is more and more is less in brief motivational interventions: Characteristics of intervention content and their associations with drinking outcomes. *Psychology of Addictive Behaviors*, 28, 1026–1040. <https://doi.org/10.1037/a0036593>
- Reid, A. E., & Carey, K. B. (2015). Interventions to reduce college student drinking: State of the evidence for mechanisms of behavior change. *Clinical Psychology Review*, 40, 213–224. <https://doi.org/10.1016/j.cpr.2015.06.006>
- Riley, R. D., Lambert, P. C., & Abo-Zaid, G. (2010). Meta-analysis of individual participant data: Rationale, conduct, and reporting. *BMJ*, 340, e221. <https://doi.org/10.1136/bmj.c221>
- Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(1), 1–36. <https://doi.org/10.18637/jss.v048.i02>
- Saunders, J. B., Aasland, O. G., Babor, T. F., de la Fuente, J. R., & Grant, M. (1993). Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption-II. *Addiction*, 88, 791–804. <https://doi.org/10.1111/j.1360-0443.1993.tb02093.x>
- Schomaker, M., & Heumann, C. (2018). Bootstrap inference when using multiple imputation. *Statistics in Medicine*, 37, 2252–2266. <https://doi.org/10.1002/sim.7654>
- Skinner, H. A., & Allen, B. A. (1982). Alcohol dependence syndrome: Measurement and validation. *Journal of Abnormal Psychology*, 91, 199–209. <https://doi.org/10.1037/0021-843X.91.3.199>
- Skinner, H. A., & Horn, J. L. (1984). *Alcohol Dependence Scale (ADS): User's guide*. Toronto, Canada: Addiction Research Foundation.
- Valente, M. J., & MacKinnon, D. P. (2017). Comparing models of change to estimate the mediated effect in the pretest-posttest control group design. *Structural Equation Modeling*, 24, 428–450. <https://doi.org/10.1080/10705511.2016.1274657>
- van Zundert, C. H. J., & Miočević, M. (2020). A comparison of meta-methods for synthesizing indirect effects. *Research Synthesis Methods*, 11, 849–865. <https://doi.org/10.1002/jrsm.1445>
- *Walters, S. T., Vader, A. M., Harris, T. R., Field, C. A., & Jouriles, E. N. (2009). Dismantling motivational interviewing and feedback for college drinkers: A randomized clinical trial. *Journal of Consulting and Clinical Psychology*, 77, 64–73. <https://doi.org/10.1037/a0014472>
- White, H. R., & Labouvie, E. W. (1989). Towards the assessment of adolescent problem drinking. *Journal of Studies on Alcohol*, 50, 30–37.
- *White, H. R., Mun, E. -Y., & Morgan, T. J. (2008). Do brief personalized feedback interventions work for mandated students or is it just getting caught that works? *Psychology of Addictive Behaviors*, 22, 107–116. <https://doi.org/10.1037/0893-164X.22.1.107>
- Wilson, S. J., Polanin, J. R., & Lipsey, M. W. (2016). Fitting meta-analytic structural equation models with complex datasets. *Research Synthesis Methods*, 7, 121–139. <https://doi.org/10.1002/jrsm.1199>
- *Wood, M. D., Capone, C., Laforge, R., Erickson, D. J., & Brand, N. H. (2007). Brief motivational intervention and alcohol expectancy challenge with heavy drinking college students: A randomized factorial study. *Addictive Behaviors*, 32, 2509–2528. <https://doi.org/10.1016/j.addbeh.2007.06.018>
- *Wood, M. D., Fairlie, A. M., Fernandez, A. C., Borsari, B., Capone, C., Laforge, R., & Carmona-Barros, R. (2010). Brief motivational and parent interventions for college students: A randomized factorial study. *Journal of Consulting and Clinical Psychology*, 78, 349–361. <https://doi.org/10.1037/a0019166>
- Wu, J. -Y., & Kwok, O. (2012). Using SEM to analyze complex survey data: A comparison between design-based single-level and model-based multilevel approaches. *Structural Equation Modeling*, 19, 16–35. <https://doi.org/10.1080/10705511.2012.634703>
- Zeger, S. L., Liang, K. -Y., & Albert, P. S. (1988). Models for longitudinal data: A generalized estimating equation approach. *Biometrics*, 44, 1049–1060. <https://doi.org/10.2307/2531734>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.