Journal of Consulting and Clinical Psychology
2016, Vol. 84, No. 3, 221–229

0022-006X/16/$12.00 http://dx.doi.org/10.1037/ccp0000074

Therapist Empathy, Combined Behavioral Intervention, and Alcohol Outcomes in the COMBINE Research Project

Theresa B. Moyers and Jon Houck
University of New Mexico

Samara L. Rice
University of New Mexico and State University of New York at Buffalo

Richard Longabaugh
Brown University

William R. Miller
University of New Mexico

Objective: Common factors such as therapist empathy play an important role in treatment for addictive behaviors. The present study was a secondary analysis designed to evaluate the relation between therapist empathy and alcohol treatment outcomes in data from a large, multisite, randomized controlled trial. ***Method:*** Audio-recorded psychotherapy sessions for 38 therapists and 700 clients had been randomly selected for fidelity coding from the combined behavioral intervention condition of Project COMBINE. Sessions were evaluated by objective raters for both specific content (coping with craving, building social skills, and managing negative mood) and relational components (empathy level of the therapist). Multilevel modeling with clients nested within therapists evaluated drinks per week at the end of treatment. ***Results:*** Approximately 11% of the variance in drinking was accounted for by therapists. A within-therapist effect of empathy was detected ($B = -0.381$, $SE = 0.103$, $p < .001$); more empathy than usual was associated with subsequent decreased drinking. The Social and Recreational Counseling module ($B = -0.412$, $SE = 0.124$, $p < .001$), Coping with Cravings and Urges module ($B = -0.362$, $SE = 0.134$, $p < .01$), and the Mood Management module ($B = -0.403$, $SE = 0.138$, $p < .01$) were also associated with decreased drinking. No between-therapist effect was detected, and the Empathy × Module Content interactions were not significant. ***Conclusions:*** The results of the study appear consistent with the hypothesis that skills building and therapist empathy are independent contributions to the overall benefit derived from the combined behavioral intervention.

What is the public health significance of this article?
This study suggests that the interpersonal skills of the therapist influence the effectiveness of a behavioral treatment for problem drinking.

Keywords: combine, empathy, therapist effects, skills-building, alcohol

Two decades of randomized controlled trials (RCTs) have yielded conclusive evidence that psychosocial treatments are a worthwhile addition to the array of interventions now available for problem drinking. Despite the clear advantage of these treatments for a person seeking to change problematic drinking, there is still little evidence to help us understand how they convey the advantage they do. Efforts to explore specific elements and theory-driven procedures as causal mechanisms in empirically supported treatments (ESTs) for addiction have often failed to support the theories generating them (Bergmark, 2008; Morgenstern & McKay, 2007; Magill & Longabaugh, 2013). An alternative perspective is that the value of these treatments derives from factors common to them, such as engaging the client's hope and providing an acceptable rationale for change (Anderson, Lunnen, & Ogles, 2010; Bohart & Wade, 2013). Characteristics of therapists who deliver these treatments are sometimes cited as a possible ingre-

This article was published Online First January 21, 2016.

Theresa B. Moyers, Psychology Department, University of New Mexico; Jon Houck, Center on Alcoholism, Substance Abuse and Addictions, University of New Mexico; Samara L. Rice, Center on Alcoholism, Substance Abuse and Addictions, University of New Mexico, and Research Institute on Addictions, State University of New York at Buffalo; Richard Longabaugh, Psychiatry and Human Behavior, Brown University Medical School; and William R. Miller, Psychology Department, University of New Mexico.

Supported by the National Institute on Alcohol Abuse and Alcoholism of the National Institutes of Health (Grants U10AA11716, K01AA021431, T32-AA018108, and T32-AA007583). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We thank Dr. Joseph Lucke for statistical support in preparation of this article.

Correspondence concerning this article should be addressed to Theresa B. Moyers, Department of Psychology MSC03 2220, University of New Mexico, Albuquerque, NM, 87131-0001. E-mail: tmoyers@unm.edu

dient in their effectiveness, particularly as they relate to the therapists' ability to form a strong alliance with the client (Anderson, Ogles, Patterson, Lambert, & Vermeersch, 2009; Crits-Christoph, et al., 2009; Wampold & Brown, 2005). While stable characteristics such as age, gender, and professional background and training have shown little association with treatment outcomes (Beutler et al., 2004), the interpersonal skills of therapists as measured during treatment sessions are more promising (Crits-Christoph, et al., 2009; Norcross & Lambert, 2011; Moyers & Miller, 2013). Therapist effects have sometimes been prominent in RCTs for alcohol use disorders (Project MATCH Research Group, 1998), but efforts to investigate the contribution of these interpersonal skills have been constrained by the rigor of the RCT methodology and by the large sample of therapists necessary to adequately examine them. Large RCTs do not typically lend themselves to careful measurement of therapist interpersonal skills; instead, they focus on whether and how the therapist delivers the content of the treatment that is the larger focus of the study. When therapist effects are subsequently detected, and there is a paucity of data to explain what might account for them, controversy concerning the relative importance of therapists in ESTs for addictions intensifies and polarizes.

An example of a rigorous RCT that avoided this particular dilemma was the COMBINE Study (Anton et al., 2006). Designed to investigate the efficacy of two medications in the treatment of heavy drinking, COMBINE also included a start-of-the art psychosocial treatment: the combined behavioral intervention (CBI). CBI is relatively unique among psychosocial treatments deriving from RCTs in that it does not rely on a single theoretical rationale. Instead, CBI blends elements from empirically supported alcohol interventions. Initial sessions focus on motivational interviewing (MI) to enhance commitment to abstinence, followed by a functional analysis to reveal specific skills deficits and high-risk situations associated with drinking (Gulliver, Longabaugh, Davidson, & Swift, 2005). Using the decision-tree model described by Longabaugh and Morgenstern (1999), skills-building modules are then selected by the therapist and client to address concerns such as management of negative emotions, responding to cravings for alcohol, and strategies for finding employment. Recruiting significant others to enrich the client's social network as well as actively facilitating attendance in mutual support groups are universal components of the treatment (Miller, 2004). CBI is therefore standardized in that it is composed of well-defined and replicable components, some of which are part of every client's treatment. It is simultaneously flexible in that it incorporates client choice from a menu of options for a large portion of the treatment, allowing a greater degree of external validity.

CBI is also relatively unusual in that it places equivalent emphasis on the *manner* in which this cognitive–behavioral content is delivered to the client. MI comprises the initial sessions, but is also strongly emphasized as the interpersonal foundation for the entire treatment (Miller, 2004; Moyers & Houck, 2011). The clinician style is explicitly empathic, collaborative, and supportive of the client's autonomy, placing relatively less value on expertise and direction on the part of the clinician than found in traditional cognitive–behavioral treatments for problem drinking.

The findings of the COMBINE Study demonstrated that CBI is effective when combined with naltrexone and when combined with a brief psychosocial intervention focused on medication adherence and naltrexone (Anton et al., 2006). The advantage of CBI was most evident in the 1-year period after treatment, with those receiving CBI being 20% more likely to have a good clinical outcome than those not receiving it (Donovan et al., 2008). Additional evidence has suggested that CBI is particularly useful for clients whose drinking goal was not complete abstinence, despite the fact that abstinence is the stated goal of the CBI treatment (Bujarski, O'Malley, Lunny, & Ray, 2013).

The negotiated modules in CBI address common skill deficits encountered in treating problem drinkers. Examples of these modules include management of negative emotions, coping with cravings for alcohol, and strategies for finding employment, all of which are typical of the content found in effective interventions for problem drinking such as the community reinforcement approach (Miller, Forcehimes, & Zweben, 2011) and cognitive–behavioral interventions (Mastroleo & Monti, 2013). The acquisition of new skills is therefore a potential mechanism for explaining the value of CBI to the drinkers in the COMBINE Study. In addition, the interpersonal context of the intervention bears examination as a possible active mechanism in CBI, because it was an explicit and rigorously trained component.

During the COMBINE Study, CBI therapy tapes were randomly selected and reviewed to assess for fidelity to both the explicit content of the modules (skills-building) and the manner in which the clinician delivered that explicit content (interpersonal skills of the clinician; Miller, Moyers, Arciniega, Ernst, & Forcehimes, 2005). For the explicit content of the modules, a checklist was used to ensure that clinicians included important elements (e.g., rehearsing drink refusal skills). An additional rating of that same session evaluated clinician skill in five areas: direction, empathy, MI style, protocol adherence, and interpersonal skills (warmth, egalitarianism, and acceptance). These characteristics were rated on a 7-point Likert-type scale after the rater listened to the entire CBI session. Empathy was operationalized as the ability of the clinician to convey an understanding of the client's perspective. Behaviorally, this was defined as (a) accurate reflective listening, (b) questions that indicated an understanding of the client's discourse, and (c) explicit recognition of the client's unspoken emotions or meaning. Factors such as warmth, collaboration, and support of autonomy were explicitly excluded from the empathy rating in favor of a definition consistent with previous research in client-centered therapy (Zuroff, Kelly, Leybman, Blatt, & Wampold, 2010.

This approach to collecting quality assurance measures for both the content and the process of CBI during the COMBINE Study yielded an opportunity to evaluate the presence of two very different hypothesized mechanisms of action in a bona fide addiction treatment and to explore their relation to drinking outcomes. Such an approach addressed the controversy concerning the contribution of clinicians to client outcomes. Stated in the most polarized manner, the controversy concerns whether general contextual factors, such as therapist characteristics, are responsible for the value clients gain from psychosocial interventions (Wampold & Brown, 2005) or whether it is specific, theory-driven elements that ought to be the focus of scientific attention (Baker, McFall, & Shoham, 2008; Chambless & Ollendick, 2001). Of course, the polemic nature of this question overlooks the possibility that specific elements *and* the interpersonal skills of therapist might both make substantive contributions to improved client outcomes (Miller & Moyers, 2015). Research to explore such a mutual contribution

would require measurement of both theory-driven elements of treatment and the interpersonal skills of the therapist delivering them.

This study explored the contribution of specific treatment modules from CBI sessions, as well as the interpersonal skill of the clinician offering them, and the relation of both to drinking outcomes in a large, multisite RCT of addictions treatment. Because therapist empathy is a specific, hypothesized active ingredient in empirically based relationships (Norcross & Wampold, 2011), and specifically within addictions treatments (Moyers & Miller, 2013), we predicted that therapists with higher empathy ratings would have clients who drank less at the end of treatment. We also hypothesized that receiving treatment modules would be associated with better drinking outcomes, and we predicted Therapist Empathy × Module Content interaction such that clients receiving modules delivered with higher therapist empathy would show optimal outcomes.

Method

Overview of Study Design

Project COMBINE was a randomized, placebo-controlled study of two pharmacotherapies, naltrexone and acamprosate, and two manualized behavioral interventions, the CBI (Miller, 2004) and medical management (MM: Pettinati et al., 2005). Eleven sites recruited 1,383 participants for 16 weeks of active treatment. Detailed description of study participants is available elsewhere (Anton et al., 2006). All procedures for this study were approved and overseen by each site's institutional review board, and written consent was obtained prior to participation. Participants were assigned to one of nine treatment cells: MM with placebo, MM with acamprosate, MM with naltrexone, MM with acamprosate plus naltrexone, CBI with placebo, CBI with acamprosate, CBI with naltrexone, CBI with acamprosate plus naltrexone, and CBI with no medication.

Assessment of participants was conducted at baseline, with postrandomization follow-up assessments occurring at weeks 8, 16, 26, 52, and 68. Assessments yielded measures related to drinking behavior: percent days abstinent (PDA), drinks per drinking day (DDD), and a composite clinical outcome measure that integrated both client alcohol consumption and alcohol-related problems (Cisler & Zweben, 1999). A complete description of measures used in the COMBINE Study can be found elsewhere (Anton et al., 2006).

The present study focused on clients randomized to CBI. Data from participants were analyzed regardless of their later compliance with or discontinuation of their assigned treatment.

CBI Therapist Selection, Training, and Supervision

Demographic data for therapists in the COMBINE Study were available only for the 31 of 38 who agreed to complete an optional assessment packet. Demographic data for these 31 therapists (81%) are found in Table 1. All 38 therapists treated an average of 18.42 clients (SD = 14.05; range: 1−47 clients).

The Training and Quality Assurance Center (TQAC) for CBI was administered by the Center on Alcoholism, Substance Abuse and Addictions at the University of New Mexico. Therapists eligible to participate in the CBI condition were required to have at least a master's degree and the appropriate license in psychology, counseling, social work, or a related behavioral health field. In addition, they were required to have at least two years of postdegree counseling experience. Candidates who met these criteria were required to submit audiotapes of two 10-min practice sessions to demonstrate their proficiency in empathy. These sessions were rated on dimensions related to empathy in the Motivational Interviewing Skills Code (Miller, 2000) and 57 therapists met the standards necessary to be accepted for further training (Miller et al., 2005).

Therapists then completed workshop training for the CBI (Miller, 2004) and viewed MI training videos. Therapists who had passed the empathy prescreen were enrolled in a training workshop.

After initial training, all CBI therapists submitted complete audio recordings for at least two practice cases that met protocol standards, using either pilot clients or role-plays scripted by the TQAC. If certified, the therapist began treating study participants. If performance did not meet protocol standards, additional cases were required prior to approval. Following certification, therapists were required to tape-record every CBI session with every client. A total of 939 sessions were reviewed by the TQAC during the study. For this project, we excluded role-played sessions (n = 35) resulting in 904 sessions with actual clients, constituting 11. 8% of the 7,674 total CBI sessions conducted in the trial (Miller et al., 2005).

Table 1
Therapist Descriptive Statistics (N = 38)

Variable	*f/M*	*SD*
Sex		
Male	19	
Female	12	
Missing	7	
Marital status		
Single	6	
Cohabitating	3	
Divorced	5	
Married	16	
Engaged	1	
Missing	7	
Ethnicity		
Hispanic	4	
Non-Hispanic	27	
Missing	7	
Degree		
Master's	15	
PhD/PsyD	16	
Missing	7	
Experience (months)		
Total clinical	100.85	94.81
Substance specific	74.20	73.38
Patient contact (hours per week)		
Substance patients	14.10	11.73
All patients	6.08	8.27

Note. *f/M* reports the frequency for categorical variables and mean for continuous variables.

Raters

Raters for the project were six graduate students at the University of New Mexico, two of whom coded 79% of the tapes. To estimate interrater reliability for empathy ratings, we calculated

intraclass correlations (ICC) using Model 3,1 (Shrout & Fleiss, 1979) from a sample of 114 double-coded sessions. Interrater reliability for the two primary coders in the study was good (ICC = .661, n = 57 sessions). Reliability for occasional coder pairs was also good (ICC = .737, n = 7; ICC = .641, n = 10).

All sessions submitted for certification or randomly selected for monitoring were coded on six therapist dimensions. All ratings were on 7-point Likert-type scales, ranging from 1 (*absence of this characteristic*) to 7 (*high levels of this characteristic*). A rating of 4 or lower for any therapist was considered failing and required remediation in consultation with the TQAC, the site Project Coordinator, and the therapist. Both the therapists and their site supervisors received copies of all ratings and feedback.

Therapist ratings included nonspecific factors (acceptance, egalitarianism, and warmth), direction, empathy, MI style, protocol adherence, and overall. These were global ratings, encompassing therapist performance throughout the session. Empathy ratings yielded a mean of 5.9 (range: 4–7).

Data Analyses

Deriving the sample. We included all available sessions from CBI that had been randomly selected and reviewed for quality assurance. This yielded 904 rated sessions and 724 clients. We excluded sessions from clients who had more than 15% of their total sessions from multiple therapists (n = 24 clients; n = 41 sessions). The remaining sample included 700 clients, 38 therapists, and 863 sessions.

Multilevel modeling. Multilevel modeling (MLM; Kreft & de Leeuw, 1998; Raudenbush & Bryk, 2002) was used to assess the association of the global rating of therapist empathy with drinks per week (DW) at end of treatment while accounting for the fact that clients were nested within therapists. All MLM was done using Hierarchical Linear Modeling software, Version 7 (Raudenbush, Bryk, & Congdon, 2004).

Empathy was chosen a priori as the primary predictor variable assessing therapist impact and was measured separately for each session. Empathy was modeled both as a characteristic of the interaction between a given client and clinician (i.e., within-therapist association) and as an attribute of the clinician (i.e., between-therapist association). For cases where empathy ratings existed on multiple sessions for the same client−clinician pair, we used the mean empathy rating across those sessions.

Of the nine optional treatment modules that could be completed by participants, three were selected for analysis. The Coping with Cravings and Urges (CRAV) module was designed to help clients discover, predict, and control situations that may leave them vulnerable to heavy drinking. Mood Management (MOOD) training was created to help clients replace drinking as a means of coping with their negative mood states. Social and Recreational Counseling (SARC) was intended to help clients identify and engage in rewarding activities that did not involve drinking. These modules were selected because they represented core aspects of theoretical interest in implementing a cognitive–behavioral approach to the cessation of heavy drinking and because they were chosen with the highest frequency by clients and their therapists (see Table 2). Further examination of treatment integrity for these modules was not possible because clinicians only reported whether they had used the module on session checklists. In addition, these particular CBI modules were not necessarily conducted during the sessions we reviewed for empathy.

Table 2
Module Completion Frequency (Client N = 700)

Module	*f*	%
Coping with craving and urges	419	59.86
Mood management training	356	50.86
Social and recreational counseling	203	29.00
Drink refusal	169	24.14
Assertion skills training	154	22.00
Mutual support group facilitation	128	18.29
Communication skills training	117	16.71
Social support for sobriety	59	8.43
Job finding training	22	3.14

Note. Due to multiple selections, percentages will not total to 100%.

Mean DW at the end of treatment was used as the primary outcome measure for our analyses. The end of the treatment time point was selected because we reasoned that the signal for both therapist empathy and specific content of modules could be detected at this point, as opposed to earlier (during treatment) or later (1 year after treatment). We used DW as our outcome variable because it encompasses both quantity and frequency in a single variable and has been used in previous studies from our research group (Moyers, Martin, Houck, Christopher, & Tonigan, 2009).

Average DW was derived by using the following formula: DDD × ([1 − PDA] × 7). DW was a count variable that ranged from 0−238, with a zero count (i.e., no DW) for 37.6% of the scores at the end of treatment. A Poisson distribution for constant exposure, accounting for dispersion, was used in the multilevel model due to the non-normal distribution of the outcome variable.

Thus, a two-level MLM assessed the association of therapist empathy and the CRAV, SARC, and MOOD modules with DW at end of treatment. Empathy was group-mean centered for the within-therapist association, grand-mean centered for the between-therapist association, and the modules were coded −0.5 if the client did not select that module or +0.5 if they did. Restricted predictive quasi-likelihood was used to estimate parameters, and the unit-specific model with robust standard errors was interpreted. Because Poisson regression raw regression coefficients are on a log scale, they are typically exponentiated and interpreted as rate ratios (Atkins, Baldwin, Zheng, Gallop, & Neighbors, 2013). Rate ratios are interpreted similarly to odds ratios in logistic regression, that is, the distance above or below 1 is interpreted as the percentage increase or decrease in the outcome for a 1-unit increase in the predictor.

A three-level MLM, which would have accounted for site at the third level, was considered but not specified due to concerns about lack of power and convergence problems, because there were too few sites (n = 11) to permit interpretation of clustering (Kreft & de Leeuw, 1998).

Finally, the ICC was calculated from a one-way analysis of variance source table with therapist specified as a fixed effect, as recommended by Baldwin et al. (2011). This method was chosen to allow for the possibility of a negative ICC. Calculating the ICC using the variance components estimated by HLM may have led to a biased ICC because the Hierarchical Linear Modeling software constrains variance estimates to be positive (nonnegativity con-

straint; Baldwin et al., 2011; Swallow & Monahan, 1984). In addition, because a Poisson distribution was specified for the outcome variable, the Level 1 equation did not have a lowest level error term associated with it. This error term is required to compute an ICC (Hox, 2010), so we, instead, calculated the ICC using the normal distribution with two different versions of the outcome variable: an untransformed and a log-transformed version.

Results

Therapist Empathy and Module Content as Predictors of Drinking Outcomes

Proportion of drinking outcomes accounted for by therapist. The ICC describing the relation between the particular therapist assigned to the client and end-of-treatment drinking was .214, 95% confidence interval (CI) [.108, .338], when calculated with the untransformed outcome variable, and .114, 95% CI [.029, .221], when calculated as a log-transformed variable. Because many participants were not drinking during the last week of treatment, a better estimate of the ICC is that approximately 11% of the variance in log-transformed DW at end of treatment was accounted for at the therapist level.

Association between therapist empathy and client drinking. At Level 1, within-therapists, the relation of therapist empathy with drinking was in the predicted direction: Therapist empathy was inversely associated with client drinking at the end of treatment ($B = -0.381$, $SE = 0.103$, $p < .001$). That is, when therapists expressed more empathy than they usually did, the client drank less at the end of treatment, and, conversely, when therapists expressed less empathy than they usually did, the client drank more at the end of treatment.

However, at Level 2 (between-therapists), the test of the between-therapist association of empathy on client DW was nonsignificant ($B = -0.071$, $SE = 0.244$, $p = .772$). That is, after statistically adjusting for the association of within-therapist empathy on client drinking, the average level of empathy between therapists (whether that therapist expressed more or less empathy in general than other therapists in the study) was not predictive of client drinking.

Variables conventionally used to test for therapist effects, such as age, gender, client-therapist gender or ethnicity match, months of general clinical experience, as well as months specifically treating substance abuse clients, were tested but were not significantly associated with client drinking, and were thus removed from the MLM model. Similarly, variables conventionally used to test for client effects, such as readiness to change and working alliance, were also tested but not significantly associated with client drinking, and were also removed from the MLM.

Association between module completion and client drinking. Modules were included as fixed effects; the frequencies of module completion can be found in Table 2. All three modules were significantly associated with drinking at the end of treatment: CRAV ($B = -0.362$, $SE = 0.134$ $p = .01$), MOOD ($B = -0.403$, $SE = 0.138$, $p < .01$), and SARC ($B = -0.418$, $SE = 0.147$, $p < .01$).

Interaction of therapist empathy with CBI module completion. There were no significant Empathy × Module Selection interactions—knowing the level of empathy provided no information on the association of the module with drinking outcomes. Each interaction was tested separately and also combined in a single model. Because there were no significant interactions, all were dropped from the final model (see Table 3).

Discussion

In the CBI, the particular therapist assigned to a client accounted for 11% of the variance in end-of-treatment drinking in the delivery of a highly structured, manualized, evidence-based treatment. This finding is consistent with a growing body of research that has indicated therapists are not interchangeable and are likely to bring personal characteristics and behaviors to the treatment process related to better and worse outcomes for their clients.

We had hypothesized that differences in empathic skills of therapists would be associated with better drinking outcomes, perhaps accounting for some of the overall impact of therapists in

Table 3
Overdispersed Poisson Multilevel Model of the Association of Therapist Empathy and CBI Modules on Drinks per Week at the End of Treatment

Fixed	Intercept coefficient	*SE*	*t* ratio	*df*	*p*	Rate ratio
Intercept	2.431	0.112	21.655	36	.001	11.370
Empathy−2[a]	−0.071	0.244	−0.292	36	.772	0.931
Empathy−1[b]	−0.381	0.103	−3.688	617	.001	0.683
CRAV module	−0.362	0.134	−2.691	617	.01	0.696
MOOD module	−0.403	0.138	−2.918	617	.01	0.668
SARC module	−0.418	0.147	−2.847	617	.01	0.658

Random	Variance component	*SD*	χ^2	*df*	*p*
Intercept	0.285	0.534	107.616	36	.001
Level 1	31.702	5.630			

Note. Unit-specific model with robust standard errors. CRAV = Coping with Craving and Urges; MOOD = Mood Management; SARC = Social and Recreational Counseling.
[a] Between-therapist empathy modeled at Level 2. [b] Within-therapist empathy modeled at Level 1.

this treatment. This hypothesis was not supported. Instead, we found that therapists were similar in their overall level of empathy compared with each other but were different in their level of empathy within their own client pool. This variability in the empathic relationship among clients *was* associated with outcomes, such that relatively small increases in the therapist's usual level of empathy were associated with larger decreases in end-of-treatment drinking. That we were able to see a relation between empathy and drinking outcomes in this large RCT is noteworthy because therapists were rigorously screened for their use of empathic listening skills prior to hiring, and were extensively and explicitly trained in the interpersonal context of the CBI (Miller et al., 2005). Further, therapists were monitored in their expression of empathy as the trial progressed and were red-lined (i.e., stopped from taking clients) if empathy ratings were unacceptably low at any point. Such procedures likely contributed to an emphasis on and consistency in the expression of empathy in CBI sessions, yielding a restricted range in ratings.

In MLM analyses, between-therapist effects ought to be seen if clinicians vary from each other in their ability to convey empathy, and those differences might then be related to a meaningful outcome such as drinking. Within-therapist effects, on the other hand, ought to be seen if clinicians vary among their own clients in their ability to convey empathy, and would indicate more to do with how well each therapist and client pair work together. Both within- and between-therapist effects are candidates for influencing the delivery of ESTs, although within-therapist effects have the potential to explain a larger percentage of the variance in client outcomes than the typically modest impact of small between-therapist effects (Baldwin & Imel, 2013).

In our data, we observed within-therapist effects, but the between-therapist effects were not significant. This indicates that the level of empathy in the CBI treatment does not reside solely within the therapist, but is more properly thought of as an interactive phenomenon similar to working alliance. If true, this could imply a smaller burden on the selection and training of therapists toward attaining higher levels of interpersonal skills and a relatively greater burden on the best match between therapist and client characteristics toward a better facilitation of empathy (Lambert & Baldwin, 2009). Such a conclusion should be approached with caution, however. Because large differences in interpersonal skills (including empathy) are often detected for therapists in less constrained settings (Anderson et al., 2009; Zuroff et al., 2010), including substance use disorder treatments (Crits-Christoph et al., 2009; Moyers & Miller, 2013), it is possible that the lack of between-therapist effects for empathy seen here is an artifact of this particular study design. In any case, our data indicate that differences in the way therapists express empathy in their interactions with individual clients might be just as important as their overall level of empathy compared with other therapists.

When a counselor's expressed empathy is related to differences among clients but not between them, it is not possible to rule out a confounding variable as the explanation for the observed association between empathy and drinking outcomes. For example, it could be that therapists expressed higher empathy with clients who were more motivated, and that those more motivated clients were also more likely to change their drinking. In other words, our a priori variable—clinician empathy—might be a coincidental marker for some other client characteristic that is closely related to the client's improved drinking. This type of potential confound cannot be ruled out with correlational data of the kind reported here. Experimental studies would be needed to address the causal value of empathy more confidently.

Nevertheless, our data are consistent with a growing body of literature supporting the hypothesis that therapists' ability to construct relationships with clients can be a key factor in determining treatment outcomes (Baldwin, Wampold, & Imel, 2007; Kraus, Castonguay, Boswell, Nordberg, & Hayes, 2011; Norcross & Lambert, 2011). This is particularly apparent with therapist-centered MLM approaches and when the sample of therapists is adequate (Baldwin & Imel, 2013; de Jong, Moerbeek, & van der Leeden, 2010; Zuroff et al., 2010). Further, this study extends such findings from mental health problems, such as depression and panic disorder, to the arena of problem drinking and supports the call for additional research into the complex interpersonal processes that comprise addiction treatments (Bergmark, 2008; Knuuttila, Kuusisto, & Saarnio, 2011; Longabaugh et al., 2005).

We found that the content of all the cognitive–behavioral modules we examined was independently associated with improvement in drinking. Teaching skills to improve social and recreational domains of a client's life, coping with cravings for alcohol, and managing negative emotions predicted better outcomes regardless of the therapist's level of empathy. We were able to detect this effect despite the fact that the measurement of the particular modules we examined was extremely crude, simply indicating whether it was present or absent. Because we only examined the three most common modules used, and because we focused on end-of-treatment drinking (when the impact of skills-building modules might be less important than later follow-up points), emphasis on the specific value of one module over another in the CBI structure should be tentative. Further, the selection of which module was negotiated, not prescribed, so that the overall behavioral content the client received in CBI was confounded with other influences from each therapist−client pair. Nevertheless, these results indicate an advantage for using cognitive–behavioral techniques to reduce depression, cope with cravings, and manage social challenges when problem drinkers are agreeable.

The lack of a Therapist Level of Empathy × Treatment Modules interaction was surprising to us, because we had speculated that therapist empathy would facilitate the acquisition of skills from modules in addition to a direct relation with outcome. But, in some ways, this lack of an interaction is encouraging. It is consistent with the hypothesis that there are multiple, independent pathways that will be helpful when clients seek a change in their drinking from a psychosocial intervention. CBI appears to work, in part, by helping clients to build new skills to enhance their quality of life, consistent with a large body of research that has demonstrated the value of cognitive–behavioral treatment for problem drinking. These gains appear to be likely, regardless of how empathic the therapist is when delivering the module content. Therapist empathy may constitute a separate pathway for improvement, which is likewise influential, regardless of the session content. Our data support the hypothesis that both the interpersonal

environment of the intervention, and the content of it, are important indicators of treatment effectiveness.

Several weaknesses of this study limit the usefulness of our findings. In particular, the lack of competence ratings in the content of the CBI modules limits the conclusions we can draw about their impact on drinking. Although our ratings ensured that the modules were offered comprehensively by means of a checklist of the critical elements, it is possible that more fine-grained examination of therapist competence in delivering the CBI modules would show a larger association with outcome.

Another limitation concerns the lack of client ratings for the empathy variable of interest. Because client ratings of therapist empathy sometimes predict outcomes more accurately than observer ratings (Elliott, Bohart, Watson, & Greenberg, 2011), it would have been useful to have a converging measure from clients for each session regarding the interpersonal skills of the therapist. Such ratings would have been particularly useful given the within-therapist effects we found for CBI therapists.

Despite these limitations, this study has several strengths that make it relatively unique among research projects attempting to investigate therapist effects in RCTs focusing on substance use. First, our sample of 38 therapists and 700 clients permitted greater confidence in modeling the Level 2 variable (therapists) than is typically possible in smaller RCTs investigating behavioral treatments for substance misuse disorders. Second, we were able to measure empathy reliably using objective raters listening to audio recordings of therapy sessions. Third, the selection of tapes for review was done randomly and raters were masked to client outcomes, eliminating the possibility of a rater bias. Fourth, the effect seen in our data, despite the restricted range of therapist empathy expressed in this sample, suggests that our effect size is probably lower than would be seen in a more typical study with a broader range of empathy scores.

Finally, the process variable of interest in this study—therapist empathy—was originally included for measurement in the COMBINE Study because it could be operationally defined and was relevant to a large body of literature on therapist interpersonal skills in the field of psychotherapy more generally (Elliott et al., 2011) and addictions treatment more specifically (Campbell, Guydish, Le, Wells, & McCarty, 2014; Moyers & Miller, 2013). Our data indicate that some variables considered common factors can be measured and evaluated for their impact on outcome in much the same way that theory-driven, specific elements have been (Miller & Moyers, 2015). Furthermore, because empathy is in part an interpersonal skill that therapists can learn, and not only a personality characteristic they either have or do not have (Crits-Christoph et al., 2006; de Roten, Zimmermann, Ortega, & Despland, 2013; Hill & Knox, 2013; Zuroff et al., 2010), it is possible to imagine that this important clinical skill could be manipulated in a RCT to evaluate its impact in behavioral treatments for problem drinking. Our study contributes to the potential for such rigor by demonstrating that therapist empathy can be operationally defined, can be measured reliably by objective observers as it is expressed during therapy sessions, and bears a significant relation to drinking at the end of treatment. It is hard to imagine a treatment factor more specific than that.

References

Anderson, T., Lunnen, K. M., & Ogles, B. M. (2010). Putting models and techniques in context. In B. Duncan, S. D. Miller, B. E. Wampold, & M. A. Hubble (Eds.), *The heart and soul of change* (2nd ed., pp. 143–166). Washington, DC: American Psychological Association.

Anderson, T., Ogles, B. M., Patterson, C. L., Lambert, M. J., & Vermeersch, D. A. (2009). Therapist effects: Facilitative interpersonal skills as a predictor of therapist success. *Journal of Clinical Psychology, 65,* 755–768. http://dx.doi.org/10.1002/jclp.20583

Anton, R. F., O'Malley, S. S., Ciraulo, D. A., Cisler, R. A., Couper, D., Donovan, D. M., . . . COMBINE Study Research Group. (2006). Combined pharmacotherapies and behavioral interventions for alcohol dependence: The COMBINE study: A randomized controlled trial. *JAMA: Journal of the American Medical Association, 295,* 2003–2017. http://dx.doi.org/10.1001/jama.295.17.2003

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. http://dx.doi.org/10.1037/a0029508

Baker, T. B., McFall, R. M., & Shoham, V. (2008). Current status and future prospects of clinical psychology: Toward a scientifically principled approach to mental and behavioral health care. *Psychological Science in the Public Interest, 9,* 67–103.

Baldwin, S. A., & Imel, Z. E. (2013). Therapist effects. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed., pp. 258–297). Hoboken, NJ: Wiley.

Baldwin, S. A., Murray, D. M., Shadish, W. R., Pals, S. L., Holland, J. M., Abramowitz, J. S., . . . Watson, J. (2011). Intraclass correlation associated with therapists: Estimates and applications in planning psychotherapy research. *Cognitive Behaviour Therapy, 40,* 15–33. http://dx.doi.org/10.1080/16506073.2010.520731

Baldwin, S. A., Wampold, B. E., & Imel, Z. E. (2007). Untangling the alliance-outcome correlation: Exploring the relative importance of therapist and patient variability in the alliance. *Journal of Consulting and Clinical Psychology, 75,* 842–852. http://dx.doi.org/10.1037/0022-006X.75.6.842

Bergmark, A. (2008). On treatment mechanisms—What can we learn from the COMBINE study? *Addiction, 103,* 703–705. http://dx.doi.org/10.1111/j.1360-0443.2007.02073.x

Beutler, L. E., Malik, M., Alimohamed, S., Harwood, T. M., Talebi, H., Noble, S., & Wong, E. (2004). Therapist variables. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed., pp. 227–306). Hoboken, NJ: Wiley.

Bohart, A. C., & Wade, A. G. (2013). The client in psychotherapy. In M. J. Lambert (Ed.), *Bergin & Garfield's handbook of psychotherapy and behavior change* (6th ed., pp. 219–257). Hoboken, NJ: Wiley.

Bujarski, S., O'Malley, S. S., Lunny, K., & Ray, L. A. (2013). The effects of drinking goal on treatment outcome for alcoholism. *Journal of Consulting and Clinical Psychology, 81,* 13–22. http://dx.doi.org/10.1037/a0030886

Campbell, B. K., Guydish, J., Le, T., Wells, E. A., & McCarty, D. (2014). The relationship of therapeutic alliance and treatment delivery fidelity with treatment retention in a multisite trial of twelve-step facilitation. *Psychology of Addictive Behaviors, 29,* 106–113. http://dx.doi.org/10.1037/adb0000008

Chambless, D. L., & Ollendick, T. H. (2001). Empirically supported psychological interventions: Controversies and evidence. *Annual Review of Psychology, 52,* 685–716. http://dx.doi.org/10.1146/annurev.psych.52.1.685

Cisler, R. A., & Zweben, A. (1999). Development of a composite measure for assessing alcohol treatment outcome: Operationalization and validation. *Alcoholism: Clinical and Experimental Research, 23,* 263–271.

Crits-Christoph, P., Gibbons, M. B. C., Crits-Christoph, K., Narducci, J., Schamberger, M., & Gallop, R. (2006). Can therapists be trained to improve their alliances? A preliminary study of alliance-fostering psychotherapy. *Psychotherapy Research, 16,* 268–281. http://dx.doi.org/10.1080/10503300500268557

Crits-Christoph, P., Gallop, R., Temes, C. M., Woody, G., Ball, S. A., Martino, S., & Carroll, K. M. (2009). The alliance in motivational enhancement therapy and counseling as usual for substance use problems. *Journal of Consulting and Clinical Psychology, 77,* 1125–1135. http://dx.doi.org/10.1037/a0017045

de Jong, K., Moerbeek, M., & van der Leeden, R. (2010). A priori power analysis in longitudinal three-level multilevel models: An example with therapist effects. *Psychotherapy Research, 20,* 273–284. http://dx.doi.org/10.1080/10503300903376320

de Roten, Y., Zimmermann, G., Ortega, D., & Despland, J.-N. (2013). Meta-analysis of the effects of MI training on clinicians' behavior. *Journal of Substance Abuse Treatment, 45,* 155–162. http://dx.doi.org/10.1016/j.jsat.2013.02.006

Donovan, D. M., Anton, R. F., Miller, W. R., Longabaugh, R., Hosking, J. D., Youngblood, M., . . . the COMBINE Study Research Group. (2008). Combined pharmacotherapies and behavioral interventions for alcohol dependence (The COMBINE Study): Examination of posttreatment drinking outcomes. *Journal of Studies on Alcohol and Drugs, 69,* 5–13. http://dx.doi.org/10.15288/jsad.2008.69.5

Elliott, R., Bohart, A. C., Watson, J. C., & Greenberg, L. S. (2011). Empathy. In J. C. Norcross (Ed.), *Psychotherapy relationships that work: Evidence-based responsiveness* (2nd ed., pp. 132–152). New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780199737208.003.0006

Gulliver, S. B., Longabaugh, R., Davidson, D., & Swift, B. (2005). The development of a broad spectrum treatment for patients with alcohol dependence in early recovery. *Cognitive and Behavioral Practice, 12,* 53–63. http://dx.doi.org/10.1016/S1077-7229(05)80039-2

Hill, C. E., & Knox, S. (2013). Training and supervision in psychotherapy. In M. J. Lambert (Ed.), *Handbook of psychotherapy and behavior change* (6th ed., pp. 775–811). Hoboken, NJ: Wiley.

Hox, J. J. (2010). *Multilevel analysis: Techniques and applications* (2nd ed.). New York, NY: Routledge.

Knuuttila, V., Kuusisto, K., & Saarnio, P. (2011). Client characteristics and therapist style: A combined analysis of impact on retention and effectiveness in outpatient substance abuse treatment. *Nordisk Alkohol- & Narkotikatidskrift, 28,* 321–338. http://dx.doi.org/10.2478/v10199-011-0028-x

Kraus, D. R., Castonguay, L., Boswell, J. F., Nordberg, S. S., & Hayes, J. A. (2011). Therapist effectiveness: Implications for accountability and patient care. *Psychotherapy Research, 21,* 267–276. http://dx.doi.org/10.1080/10503307.2011.563249

Kreft, I. G. G., & de Leeuw, J. (1998). *Introducing multilevel modeling*. Thousand Oaks, CA: Sage.

Lambert, M. J., & Baldwin, S. A. (2009). Some observations on studying therapists instead of treatment packages. *Clinical Psychology: Science and Practice, 16,* 82–85. http://dx.doi.org/10.1111/j.1468-2850.2009.01147.x

Longabaugh, R., Donovan, D. M., Karno, M. P., McCrady, B. S., Morgenstern, J., & Tonigan, J. S. (2005). Active ingredients: How and why evidence-based alcohol behavioral treatment interventions work. *Alcoholism: Clinical and Experimental Research, 29,* 235–247. http://dx.doi.org/10.1097/01.ALC.0000153541.78005.1F

Longabaugh, R., & Morgenstern, J. (1999). Cognitive-behavioral coping-skills therapy for alcohol dependence. Current status and future directions. *Alcohol Research & Health, 23,* 78–85.

Magill, M., & Longabaugh, R. (2013). Efficacy combined with specified ingredients: A new direction for empirically supported addiction treatment. *Addiction, 108,* 874–881. http://dx.doi.org/10.1111/add.12013

Mastroleo, N. R., & Monti, P. M. (2013). Cognitive-behavioral treatment for addictions. In B. M. McCrady & E. E. Epstein (Eds.), *Addictions: A comprehensive guidebook* (pp. 391–410). New York, NY: Oxford University Press.

Miller, W. R. (2000). Motivational Interviewing Skills Code (MISC): Coder's manual (Version 1.0). Retrieved from http://casaa.unm.edu/download/misc1.pdf

Miller, W. R. (2004). *COMBINE Monograph Series: Vol. 1. Combined behavioral intervention manual: A clinical research guide for therapists treating people with alcohol abuse and dependence*. (NIH Pub. No. 04-5288). Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism.

Miller, W. R., Forcehimes, A. A., & Zweben, A. (2011). A community reinforcement approach. In W. R. Miller, A. A. Forcehimes, & A. Zweben (Eds.), *Treating addiction: A guide for professionals* (pp. 172–184). New York, NY: Guilford Press.

Miller, W. R., Moyers, T. B., Arciniega, L., Ernst, D., & Forcehimes, A. (2005). Training, supervision and quality monitoring of COMBINE's behavioral interventions. *Journal of Studies on Alcohol, 66* (Supp. 15), 188–195. http://dx.doi.org/10.15288/jsas.2005.s15.188

Miller, W. R., & Moyers, T. B. (2015). The forest and the trees: Relational and specific factors in addiction treatment. *Addiction, 110,* 401–413. http://dx.doi.org/10.1111/add.12693

Morgenstern, J., & McKay, J. R. (2007). Rethinking the paradigms that inform behavioral treatment research for substance use disorders. *Addiction, 102,* 1377–1389. http://dx.doi.org/10.1111/j.1360-0443.2007.01882.x

Moyers, T. B., & Houck, J. M. (2011). Combining motivational interviewing with cognitive-behavioral treatments for substance abuse: Lessons from the Combine Research Project. *Cognitive and Behavioral Practice, 18,* 38–45. http://dx.doi.org/10.1016/j.cbpra.2009.09.005

Moyers, T. B., Martin, T., Houck, J. M., Christopher, P. J., & Tonigan, J. S. (2009). From in-session behaviors to drinking outcomes: A causal chain for motivational interviewing. *Journal of Consulting and Clinical Psychology, 77,* 1113–1124.

Moyers, T. B., & Miller, W. R. (2013). Is low therapist empathy toxic? *Psychology of Addictive Behaviors, 27,* 878–884. http://dx.doi.org/10.1037/a0030274

Norcross, J. C., & Lambert, M. J. (2011). Evidence-based therapy relationships. In J. C. Norcross (Ed.), *Psychotherapy relationships that work: Evidence-based responsiveness* (2nd ed., pp. 3–22). New York, NY: Oxford Press. http://dx.doi.org/10.1093/acprof:oso/9780199737208.003.0001

Norcross, J. C., & Wampold, B. E. (2011). Evidence-based therapy relationships: Research conclusions and clinical practices. *Psychotherapy, 48,* 98–102. http://dx.doi.org/10.1037/a0022161

Pettinati, H. M., Weiss, R. D., Dundon, W., Miller, W. R., Donovan, D., Ernst, D. B., & Rounsaville, B. J. (2005). A structured approach to medical management: A psychosocial intervention to support pharmacotherapy in the treatment of alcohol dependence. *Journal of Studies on Alcohol and Drugs, Supp. 15,* 170–178. http://dx.doi.org/10.15288/jsas.2005.s15.170

Project MATCH Research Group. (1998). Therapist effects in three treatments for alcohol problems. *Psychotherapy Research, 8,* 455–474. http://dx.doi.org/10.1080/10503309812331332527

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). Thousand Oaks, CA: Sage.

Raudenbush, S. W., Bryk, A. S., & Congdon, R. (2004). *HLM 6*. Lincolnwood, IL: Scientific Software International.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86,* 420–428. http://dx.doi.org/10.1037/0033-2909.86.2.420

Swallow, W. H., & Monahan, J. F. (1984). Monte Carlo comparison of ANOVA, MIVQUE, REML, and ML estimators of variance components. *Technometrics, 26,* 47–57. http://dx.doi.org/10.1080/00401706.1984.10487921

Wampold, B. E., & Brown, G. S. (2005). Estimating variability in outcomes attributable to therapists: A naturalistic study of outcomes in managed care. *Journal of Consulting and Clinical Psychology, 73,* 914–923. http://dx.doi.org/10.1037/0022-006X.73.5.914

Zuroff, D. C., Kelly, A. C., Leybman, M. J., Blatt, S. J., & Wampold, B. E. (2010). Between-therapist and within-therapist differences in the quality of the therapeutic relationship: Effects on maladjustment and self-critical perfectionism. *Journal of Clinical Psychology, 66,* 681–697.

Received October 16, 2014
Revision received October 23, 2015
Accepted November 2, 2015 ■

Additional Journal Information