



# HHS Public Access

Author manuscript

*Subst Use Misuse*. Author manuscript; available in PMC 2016 February 26.

Published in final edited form as:

*Subst Use Misuse*. 2013 September ; 48(12): 1161–1173. doi:10.3109/10826084.2013.808540.

## Spirituality as a Change Mechanism in 12-Step Programs: A Replication, Extension, and Refinement

J. Scott Tonigan, Kristina N. Rynes, and Barbara S. McCrady

Center on Alcoholism, Substance Abuse, and Addictions, University of New Mexico, Albuquerque, New Mexico, USA

### Abstract

This National Institutes of Health funded study investigated spiritual growth as a change mechanism in 12-step programs. A total of 130 people, early 12-step affiliates with limited Alcoholics Anonymous (AA) histories, were recruited from 2007 to 2008 from AA, treatment, and community centers in a Southwestern city in the United States. A majority of the sample was alcohol dependent. Participants were interviewed at baseline and at 3, 6, and 9 months. Lagged General Linear Modeling analyses indicated that spiritual change as measured by the Religious Background and Behavior (RBB) self-report questionnaire were predictive of increased abstinence and decreased drinking intensity, and that the magnitude of this effect varied across different RBB scoring algorithms. Future research should address study limitations by recruiting participants with more extensive AA histories and by including assessments of commitment to, and practice of, AA prescribed activities. The study's limitations are noted.

### Keywords

spirituality; change mechanism; 12-step; Alcoholics Anonymous

## INTRODUCTION

Many clinicians working in substance user treatment<sup>1</sup> programs in the United States encourage 12-step attendance (Kelly, Yeterian, & Myers, 2008), and sufficient evidence has now accumulated to assert that 12-step referral is an evidence-based practice that helps many, but not all, substance users. Several meta-analyses and numerous prospective studies have now shown that 12-step attendance alone, in combination with, and after treatment, is predictive of reductions in drinking (Emrick, Tonigan, Montgomery, & Little, 1993; Kaskutas, Bond, & Humphreys, 2002; Kelly, Stout, Magill, Tonigan, & Pagano, 2011; Tonigan & Rice, 2010; Tonigan, Toscova, & Miller, 1996) and illicit drug use (Gossop, Stewart, & Marsden, 2007; Timko, Billow, & DeBenedetti, 2006; Timko & Sempel, 2004; Weiss et al., 2005; Witbrodt & Kaskutas, 2005; Worley et al., 2008). It is important to note that long-term investigations into the benefits of 12-step programs are relatively rare, and a

---

Address correspondence to J. Scott Tonigan, Center on Alcoholism, Substance Abuse, and Addictions, University of New Mexico, 2650 Yale S.E., Albuquerque, NM 87106, USA; jtonigan@unm.edu.

### Declaration of Interest

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the article.

majority of studies are limited to 12-month follow-up. Recent work also suggests that sustained 12-step attendance may even serve to off-set relapse to illicit drug use once alcohol use has occurred (Tonigan & Beatty, 2011). Understandably, then, many studies have sought to identify the prescribed 12-step behaviors and practices that predict increased abstinence (e.g., Kelly & Moos, 2003; Pagano, Friend, Tonigan, & Stout, 2004; Tonigan & Rice, 2010), with equal attention directed to understanding *why* these 12-step practices predict improved functioning.

Among the change mechanisms *specific* to 12-step programs, (e.g., anger, Kelly, Stout, Tonigan, Magill, & Pagano, 2010), none is more central to 12-step philosophy or stressed as much in 12-step practice as is spiritual growth. In fact, an explicit goal of 12-step work and practice is to catalyze a spiritual awakening (Alcoholics Anonymous World Services, 2001). The emphasis on a spiritual transformation is made clear in *Alcoholics Anonymous: The Big Book* (2001) in which it is stated that “the alcoholic at certain times has no effective mental defense against the first drink. Except in a few rare cases, neither he nor any other human being can provide such a defense. His defense must come from a Higher Power” (p. 43). To date, six studies have investigated the function(s) that changes in spirituality may serve in explaining 12-step related benefit. These studies have used cross-sectional (Oakes, 2008), partially lagged (Kaskutas, Turk, Bond, & Weisner, 2003; Robinson, Cranford, Webb, & Brower, 2007; Tonigan, 2003; Zemore, 2007), and fully lagged designs (Kelly et al., 2011). To obtain a clearer estimate of the relationships among Alcoholics Anonymous (AA) attendance, spirituality, and drinking that these studies have reported, we conducted a meta-analysis of their findings. Several factors facilitated this meta-analysis. First, all of these studies adopted a causal step approach to investigate the mediational effect of spirituality (Baron & Kenny, 1986), albeit executed to varying degrees. Using meta-analytic techniques, we were therefore able to calculate weighted estimates that summarized the magnitude and stability of key effects reported in these six studies (Hunter & Schmidt, 2004; Preacher & Hayes, 2004). Second, a majority of these studies used the same measures of AA participation, spirituality, and alcohol use. Most studies, for instance, used the Religious Background and Behavior questionnaire (RBB; Connors, Tonigan, & Miller, 1996) to measure spirituality, and two of them (Kelly et al., 2011; Zemore, 2007) used the same algorithm to score the RBB. Third, four of the six studies evaluated the role of spiritual

---

<sup>1</sup>Treatment can be usefully defined as a unique, planned, goal directed, temporally structured, multidimensional change process, which may be phase structured, of necessary quality, appropriateness and conditions (endogenous and exogenous), implemented under conditions of uncertainty, which is *bounded* (culture, place, time, etc.), which can be (un)successful (partially and/or totally), as well as being associated with iatrogenic harm and can be categorized into professional-based, tradition-based, mutual-help-based (AA, NA, etc.), and self-help (“natural recovery”) models. Whether or not a treatment technique is indicated or contraindicated, its selection underpinnings (theory-based, empirically based, principle of faith-based, tradition-based, budget-based, etc.) continues to be a generic and key treatment issue. In the West, with the relatively new ideology of “harm reduction” and the even newer quality of life (QOL) and “wellness” treatment-driven models, there are now new sets of goals in addition to those derived from/associated with the older tradition of abstinence-driven models. Conflict-resolution models may stimulate an additional option for intervention. Treatment is implemented in a range of environments; ambulatory as well as within institutions, which can also include controlled environments such as jails, prisons, and military camps. Treatment includes a spectrum of clinician–caregiver–patient relationships, representing various forms of decision-making traditions/models, (1) the hierarchical model in which the clinician-treatment agent makes the decision(s) and the recipient is compliant and relatively passive, (2) shared decision making, which facilitates the collaboration between clinician and client(s)/patient(s) in which both are active, and (3) the “informed model” in which the patient makes the decision(s). Within this planned change process, relatively recently in various parts of the world, active substance users who are not in “treatment,” as well as those users who are in treatment, have become social change agents, active advocates, and peer health counselors...which represent just a sampling of their new labels. There are no unique models or techniques used with substance users—of whatever types and heterogeneities—which are not also used with nonsubstance users. Editor’s note.

change during early AA affiliation. As such, aggregated estimates of effect are less likely to be confounded by maturation, time, and the changing needs of people as they move into and through community-based 12-step programs.

Table 1 displays the magnitude of the associations in the spirituality causal step model, both at the study and aggregate levels. Measures of AA participation were predictive of increased abstinence in all six studies. The magnitude of this effect,  $B_{\text{weighted}} = 0.27$ , was consistent with the effect sizes found in previous meta-analytic reports (Emrick et al., 1993; Tonigan et al., 1996). The paths between AA participation and spirituality ( $B_{\text{weighted}} = 0.22$ ) and between spirituality and increased abstinence ( $B_{\text{weighted}} = 0.14$ ) also were significant, but smaller in magnitude than the path between AA participation and increased abstinence. To test statistical mediation, most of the studies evaluated the degree to which the direct path between AA and abstinence decreased in magnitude when the mediator (spirituality) was added to the model (e.g., Oakes, 2008; Tonigan, 2003; Zemore, 2007). Only Kelly et al. (2011) used the more rigorous product of coefficients approach (MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002).

In five of the six studies, spirituality partially explained the association between AA participation and increased abstinence. For example, Kelly et al. (2011) reported that in a sample of adults receiving aftercare from inpatient alcohol user treatment, 28% of the positive effect of AA attendance (intake to 3 months) on frequency of abstinent days (months 10–12) was explained by changes in RBB scores from baseline to 9 months. They reported a smaller, but significant, effect in an outpatient treatment sample, e.g., 14%. Likewise, Zemore (2007) found that entering spirituality as a mediator of the association between AA involvement and abstinence from 9 to 12 months weakened the direct effect of AA on abstinence. Specifically, the odds of abstinence at 12 months that were associated with pre-post changes in AA involvement declined from 9.35 to 7.50 when the spirituality measure was entered into the model. Tonigan (2003) reported the only null finding of the six studies, a finding based upon a partially lagged design in which the spiritual measure and outcome were both measured at a 3-year follow-up. Here, Tonigan found that reported spiritual awakening reduced the magnitude of the direct path between AA attendance at 12 months and percent days abstinent at a 3-year follow-up by only 4% (an insignificant change). Several factors may have contributed to this null finding, but the most parsimonious explanation may be that the assessment interval between the measurement of AA attendance and spiritual awakening and drinking outcome was too long.

Of some import, the variability of the weighted effects in Table 1 provides an estimate of effect that may be obtained in future research. For example, from Table 1 it can be predicted that 68% of future studies investigating the temporal linkage between increased spirituality and abstinence within a 12-step context will obtain parameter estimates that range from  $\beta = .10$  to  $\beta = .18$  ( $\pm .14$ ,  $SD = 0.04$ ). This variability may, in part, be a measurement artifact. While in most of the six studies authors used the RBB to measure spirituality, for example, the methods they used to score this scale differed. In addition, authors' scoring methods sometimes combined items of the RBB that psychometric analyses (e.g., Connors et al., 1996) have shown to be categorically distinct. Focusing only on recent spiritual and religious (S/R) beliefs and practices, for example, Kelly and colleagues (2011) and Zemore

(2007) summed participants' seven most recent S/R practices into a single score. This resulted in a combination of items that Connors et al. (1996) found to belong to two distinct factors: God Consciousness, e.g., prayer, and Formal Religious Practices, e.g., attending religious services. Like Kelly et al. (2011) and Zemore (2007), Oakes (2008) also summed participants' recent spiritual practices to create a "Recent Faith Practice" score. In addition, Oakes (2008) also summed all lifetime items on the RBB to create a "Lifetime Faith Practice" score. In contrast, Tonigan (2003) used a single item requesting a binary response to the question, "have you had a spiritual awakening?" Aside from the concern that respondents may have interpreted "spiritual awakening" in vastly different ways this single-item response also lacked reliability. Overall, then, researchers have not agreed upon a method for scoring the RBB, and this disagreement may have contributed to the significant variability in spirituality path estimates across studies.

Not shown in Table 1, the external validity of the six studies we reviewed was inversely related to the rigorousness of the methods they used to test for mediation. In particular, the four studies with the highest external validity in which sample characteristics more closely mirrored the characteristics typical of community-based 12-step members relative to clinical trials that have narrow inclusion criteria (e.g., high ethnic diversity, comorbid psychiatric, and substance use disorders), also used partially lagged designs in which the mediator, spirituality, and substance use outcome were collected at the same time (Kaskutas et al., 2003; Oakes, 2008; Robinson et al., 2007; Zemore, 2007). In contrast, the one study that used a more desirable fully lagged longitudinal design (e.g., Kelly et al., 2011) probably had the least external validity of all six studies. Indeed, Kelly et al. acknowledged that their outpatient and aftercare samples were primarily non-Hispanic white males who were motivated to participate in a randomized clinical trial (Project MATCH Research Group, 1997; 1998). In addition, since Kelly et al. tested their hypotheses using Project MATCH participants, adults with comorbid substance use disorders were excluded from participating. Considering that poly-substance misuse is commonly reported in studies of alcohol abusers, e.g., 64% of alcoholics also use illicit drugs (Tsuang, Shapiro, Smith, & Schuckit, 1994), and about 50% of all alcoholics suffer comorbid illicit substance use disorders (Ross, 1993), Kelly et al.'s findings, while encouraging, should be viewed cautiously.

Among study authors there is relative agreement about the underlying reason for the mediating effect of spirituality. Specifically, authors have suggested that S/R beliefs and/or practices foster alternative and healthy responses to stressful and aversive cues that previously were associated with substance use. Robinson and colleagues (2007) explained their finding that spiritual practices (but not beliefs) accounted for 12-step related benefit because proactive spiritual behaviors were inconsistent with drinking. Kelly and colleagues (2011) offered a more cognitively based interpretation, namely that S/R beliefs and practices may facilitate schematic cognitive restructuring, which promotes adaptive responses to stressors and negative affect. However varied in specifics, these interpretations gain credibility given reports that AA attendance predicts increased behavioral coping (e.g., Morgenstern, Labouvie, McCrady, Kahler, & Frey, 1997), stimulus control (e.g., Laffaye, McKeller, Ilgen, & Moos, 2008), increased approach coping, and decreased avoidance coping (e.g., Kelly, Magill, & Stout, 2009; Timko, Finney, & Moos, 2005).

Is there a more parsimonious explanation for the beneficial effect of spirituality among 12-step members? Research has demonstrated that, in general, shared beliefs, values, norms, and practices play a significant role in *sustaining* social group membership (e.g., Terry, Hog, & White, 2000). It seems reasonable to assert that spirituality is a shared ideology among 12-step members that offers a framework for interpreting and ascribing meaning to situations, cues, social relationships, expectations, and mood states. From this viewpoint, one of the primary functions of spiritual growth and practice may be to strengthen participants' perceived cohesion and identification with the 12-step group. These processes may reinforce sustained 12-step meeting attendance that actually accounts for increased abstinence. This subtle yet important distinction has not yet been evaluated empirically, but it is worth noting that doing so would require the researcher to demonstrate that changes in spirituality predict later drinking improvement *after* first controlling for past and concurrent 12-step meeting attendance.

### The Purpose of This Study Was Threefold

The first aim was to evaluate whether the findings reported by Kelly et al. (2011) would generalize to a sample with high external validity. Thus, we used a prospective and fully lagged mediational design, identical to the one used by Kelly et al., to test whether changes in spiritual practices would explain the association between AA attendance and drinking outcomes. Aiding this replication effort, we administered measures of spirituality (RBB; Connors et al., 1996) and AA attendance that were the same as the measures used by Kelly and colleagues, and we employed similar covariates and analytic strategy.

The second study aim was to clarify the precise items and/or scales of spirituality in the RBB that accounted for the mediational effect. Related, we also sought to map the trajectory of changes in spirituality with 12-step attendance to identify potentially critical periods in which spirituality might or might not function to explain 12-step related benefit.

Third, we aimed to test the rival hypothesis that the association between spirituality and later increases in abstinence occurs because spirituality mobilizes participants' 12-step attendance which, in turn, accounts for their increased abstinence.

## METHOD

### Participants

A total of 130 participants were recruited by research staff from AA meetings, outpatient treatment centers, and community shelters. To control for the cumulative effect of past change histories and how these histories might influence the actions of the mechanisms under study, eligibility criteria were narrow in terms of lifetime and recent treatment and AA experiences. Prospective participants were excluded if they reported more than 16 weeks of lifetime AA attendance and/or if they reported having successfully achieved alcohol abstinence for 12 months or longer after they had first determined that alcohol use had become a problem. To be included, participants had to meet current *Diagnostic and Statistical Manual of Mental Disorders* (American Psychiatric Association [DSM-IV-TR], 2000) criteria for alcohol abuse or dependence,<sup>2</sup> have consumed alcohol in the prior 90 days

and have attended at least one AA meeting in the prior 3 months. Illicit drug abuse and dependence were not exclusion criteria but were documented. All assessments were done by trained and experienced research staff.

## Procedures

Prior to consent we administered an alcohol breathalyzer test to assess intoxication. The consent process proceeded only when participants' BAC was below 0.05. Once consented, participants completed the intake interview. Follow-up assessment was conducted at 3, 6, and 9 months. Strong preference was given to conducting follow-up interviews in-person; however, telephone interviews were used if required. Using procedures developed in Project MATCH Research Group (1997, 1998), if participants missed an interview but were successfully interviewed at a later follow-up point, the data were reconstructed for the missed interview. At 3 months, 96 participants (74%) were interviewed in-person and 1 participant was interviewed by telephone (0.07%). Twenty-one participants reconstructed the 3-month interview at a later assessment (16%) resulting in complete data at 3 months for 91% ( $n = 118$ ) of the participants. At 6 months, 95 participants completed in-person interviews (73%) and 5 participants completed telephone interviews (3.8%). Fourteen participants (10.8%) reconstructed 6-month data at the 9-month interview resulting in complete data on 87.7% ( $n = 114$ ) of the participants for the 6-month interview. At the 9-month interview, 86.9% ( $n = 113$ ) of the participants completed interviews, 106 (93.8%) of which were in-person and 7 (6.2%) of which were over the telephone. This was a longitudinal assessment-only study and, as such, no intervention was offered. Nevertheless, referral to treatment could be made when requested by the participant or when it was deemed warranted by clinical staff at the Center on Alcoholism, Substance Abuse, and Addictions. Participants were reimbursed \$50 for each completed assessment. All procedures and assessments were approved by the institutional review board at the University of New Mexico (UNM Protocol # 27147).

## Assessments

Participants took approximately 2 hours to complete the baseline interview, which included three semistructured interviews, 12 different self-report questionnaires, and a urine toxicology screen for five classes of illicit drugs. Follow-up interviews were modestly shorter in duration. Only those measures central to the aims of this study are described below.

**Substance Use**—Daily alcohol and illicit drug use data were collected using the *Form 90* semistructured interview (Miller, 1996), ideally in 90-day intervals. The *Form 90* has demonstrated good test-retest reliability for measuring self-reported drinks per drinking day (ICC = 0.71) and number of days abstinent from alcohol (ICC = 0.85; Tonigan, Miller, &

---

<sup>2</sup>The reader is reminded that DSM-IV criteria—which are based on consensus and are not evidence informed—results in a label—“abuse,” “dependence,” “substance use disorder”—and not in a diagnosis, which is the outcome of a “normed” diagnostic process by a person trained and certified to make the diagnosis, which is based on the collection of relevant information to make needed decisions, and which, as a minimum, should supply three basic pieces of information: (1) etiology, (2) process, and (3) prognosis. The three aforementioned terms, from such a perspective, are medicalized, pathologized, labels, and not diagnosis. Every diagnosis is indeed a classification...a nosology...a label; every label is not a diagnosis. Editor's note.



Brown, 1997). Westerberg, Miller, & Tonigan (1998) also demonstrated that the Form 90 is a reliable indicator of self-reported use of marijuana (ICC = 0.79), cocaine (ICC = 0.75), and opiates (ICC = 0.80). Three outcome measures of substance use were computed from *Form 90* data in the current study: proportion of days abstinent from alcohol (PDA), drinks per drinking day (DPDD), and proportion of days participants took their most frequently used illicit drug (IDD). Proportion of days abstinent was defined as the number of days abstinent from alcohol divided by the total number of days in the assessment period. Drinks per drinking day were defined as the number of drinks consumed per drinking day divided by the number of drinking days in a period (abstinent days not included in the denominator). Proportion of illicit drug use days was calculated by summing the number of days participants took their most frequently used drug and dividing this sum by the number of days in the assessment period. Thus, if from baseline to 3 months, a participant used cocaine more than any other illicit drug, cocaine use days were divided by the number of days in the measurement period to attain a measure of IDD for that participant (cf. Tonigan & Beatty, 2011). Measuring IDD in this way provides a conservative estimate of drug use days when a participant uses multiple drugs within one assessment period. *Urine toxicology screens* were collected at the intake and 9-month interviews using the iCassette Drug Screen-4 Panel Test (i-DOA-1145; Instant Technologies, Inc., Norfolk, VA). This instrument detected participants' recent use of cocaine, marijuana, opiates, methamphetamine, and sedatives. If a participant had a positive drug screen but reported no illicit drug use, we evaluated if the positive screen was the result of prescribed medication (e.g., methadone) by comparing self-reported use of prescribed medical and psychiatric medications on the *Form 90* to the drug screen results.

**Help-Seeking Behaviors**—The *Form 90* was used to measure the frequency with which participants received formal and informal health care and substance use treatment in an assessment period. Treatment alliance and engagement were not assessed. Number of days that participants received treatment for alcohol, drug, and emotional problems (inpatient, outpatient, or both) was summed and divided by the total number of days in the assessment period to derive a proportion of days that participants received formal treatment in an assessment period. Data from the *Form 90* interview were also used to measure the proportion of days that participants attended 12-step meetings. To calculate this proportion, the total number of 12-step attendance days reported was divided by the total number of days in an assessment window. Using the *Form 90* to calculate proportion of days in which participants receive or participate in formal and informal help seeking has been shown to be an effective and sound method to summarize help-seeking behaviors, especially when the actual number of days in an assessment window varies across individuals (Tonigan & Beatty, 2011; Tonigan, Miller, & Brown, 1997; Westerberg et al., 1998).

**Spirituality**—The RBB self-report questionnaire measures both spiritual and religious behaviors and practices (Connors et al., 1996). The RBB does not ask respondent to identify their religious denomination. At baseline interviews, this 13-item inventory evaluates both lifetime (6 items) and past year S/R practices (7 items). In follow-up interviews, the RBB is modified to include only 7 items that assess participants' S/R beliefs and practices in the past 3 months. The first item of the RBB asks participants to indicate their current S/R

status. This item is coded “0” for atheist, “1” for agnostic, “2” for unsure, “3” for spiritual, and “4” for religious. The next 6 items of the RBB ask participants to rate the frequency with which they have had a certain S/R experience or done a certain S/R activity in the last 3 months. Participants use a 1–8 scale from 1 = never to 8 = once per day or more to rate these 7 items. The baseline version of the RBB includes additional items that ask participants to rate six S/R practices they have engaged in over their lifetime using 1 = never, 2 = in the past but not now, and 3 = continued practice of the behavior. Exploratory factor analyses have shown that the RBB has two underlying factors, God Consciousness and Formal Religious Practices, both of which have excellent test–retest stability,  $r = .94$  and  $.96$ , and high internal item consistency,  $\alpha = .76$  and  $\alpha = .81$ , respectively (Connors et al., 1996).

Several alternative RBB scoring methods were used in this study; some prospectively chosen and two introduced via post hoc analyses. First, we created a global score to reflect all of participants’ recent S/R beliefs and practices. This score was identical to the score Kelly et al. (2011) and Zemore (2007) used in their studies. We calculated this score by summing ratings of the first 7 items of the RBB. Second, we created a Formal Religious Practices scale and a God Consciousness scale from the RBB. These were identical to the scales created by Connors et al. (1996). The Formal Religious Practices scale measured the extent to which participants reported praying, attending religious services, and reading or studying holy writings both throughout their lifetime and in the past year (at baseline) and throughout the past 3 months (at follow-up assessments). The God Consciousness scale measured the extent to which participants thought about God, meditated, and had direct experiences of God. The third method we used to score the RBB involved creating a modified God Consciousness scale (MGC) and a modified Religious Practices scale. We calculated these scales by shifting participants’ ratings of their meditation practice from Connors et al.’s Formal Religious Practices scale to our MGC scale.

### Statistical Methods

To replicate Kelly et al.’s (2011) statistical procedures, we transformed each of our drinking outcome variables and our 12-step attendance independent variable. We therefore arcsine transformed PDA, square root transformed DPDD, and log transformed the proportion of 12-step attendance days. We also arcsine transformed IDD. For ease of interpretation, the untransformed value of each of the variables is included in Tables 3 and 4. To test whether spirituality mediated the association between AA attendance and drinking outcomes, we employed Baron and Kenny’s (1986) causal step model. Here, we conducted four lagged prospective analyses that tested whether: (1) AA attendance from baseline to 3 months predicted substance use in months 7–9, (2) AA attendance from baseline to 3 months predicted the 6-month RBB score (e.g., spiritual practices and beliefs in months 4–6), (3) the 6-month RBB score predicted substance use in months 7–9, and (4) adding the 6-month RBB score as a mediator significantly lessened the effect of early 12-step attendance on later drinking outcomes.

Kelly et al. (2011) entered eight covariates in each of their lagged analyses, two of which had no analog in this study—treatment assignment and Project MATCH treatment site. To replicate their strategy, we entered the following fixed covariates in all analyses: age,



gender, marital and employment status, and ethnicity. We noticed that the alcohol user treatment covariate that Kelly et al. (2011) used did not incorporate the amount of treatment received as part of Project MATCH nor did they incorporate formal treatment received after the 12 weeks of Project MATCH, but before the 15-month interview (they used number of lifetime alcohol user treatments). We thought that it was critical to control for concurrent therapy to effectively isolate the effect of spirituality and AA. We, therefore, added a lagged time-varying covariate to all analyses that accounted for the proportion of days in an assessment period that participants reported receiving either inpatient or outpatient formal treatment for alcohol, drug, and/or emotional problems.

## RESULTS

### Sample Characteristics

A total of 130 participants were recruited between 2007 and 2008; 24% ( $n = 31$ ) from AA, 51% ( $n = 66$ ) from substance user treatment and community centers for the homeless, and 25% ( $n = 33$ ) by word of mouth and advertisement. Fifty-three percent of the sample was male; 35% percent ( $n = 46$ ) were non-Hispanic white, 43% ( $n = 56$ ) were Hispanic, 15% ( $n = 20$ ) were Native American, and 6% ( $n = 8$ ) were African American, Asian, or did not report their ethnicity. On average, participants were 38.65 years old ( $SD = 9.60$ ) and ranged from 19 to 64 years of age. The majority of participants lived in a residence with their children (66.9%;  $n = 87$ ), 23 (17.7%) lived with a spouse and/or family member, and none of the participants were homeless. On average, participants reported having 12.48 ( $SD = 3.04$ ) years of formal education. Prior to recruitment participants reported using alcohol about every other day (mean PDA = 0.56;  $SD = 0.33$ ), and on drinking days they had, on average, 14.02 ( $SD = 0.23$ ) standard drinks. About 69% reported using illicit drugs in the 90-day period before recruitment and these participants reported that they used illicit drugs about 2 days per week ( $M = 0.28$ ;  $SD = 33$ ).

### Reliability of 12-Step Attendance Measure

We assessed the reliability of self-reported 12-step attendance given that this variable was essential for evaluating the study hypotheses. The number of days participants attended 12-step meetings was assessed twice in each interview, once in the Form 90 semistructured interview and a second time in the General Alcoholics Anonymous Tools of Recovery (GAATOR) self-report questionnaire. Dividing each of these reports by the number of days in an assessment period provided two independent measures of the proportion of days participants attended 12-step meetings. ICCs (formula 2.1, Shrout & Fleiss, 1979) of these paired reports at the intake, 3, 6, and 9-month interviews were 0.79, 0.91, 0.91, and 0.86, respectively. Applying Cicchetti's (1994) recommendation regarding how ICCs should be interpreted, e.g., below 0.40 = poor, 0.40 to 0.59 = fair, 0.60 to 0.74 = good, and 0.75 to 1.00 = excellent, we determined that self-reported 12-step attendance was highly reliable during follow-up.

### Reconstructed Drinking Data

A total of 345 follow-up interviews were conducted across the 3, 6, and 9-month follow-ups, and 10.1% ( $n = 35$ ) of these interviews reconstructed data for a missed 3-month and/or 6-

month interviews. No mean differences were found between reconstructed ( $n = 21$ ) and nonreconstructed ( $n = 97$ ) 3-month data on key drinking measures including PDA,  $t(116) = -0.93, p = .36$ ; DPDD,  $t(116) = 0.02, p = .98$ ; and peak BAC,  $t(116) = 0.76, p = .47$ . Likewise, there were no mean differences between reconstructed ( $n = 14$ ) and nonreconstructed drinking data for the 6-month interview on PDA,  $t(112) = 0.06, p = .95$ ; DPDD,  $t(112) = -0.75, p = .45$ ; or peak BAC,  $t(112) = -0.67, p = .51$ .

Table 2 provides descriptive data on the help seeking and substance use of participants over the course of the study. As shown, at each follow-up more than 75% of the participants reported attending 12-step meetings and a slight majority reported receiving formal treatment by professionals for their alcohol use, emotional and/or drug use-related problems. Moderate pre-post treatment substance use-related improvements were found. Proportion of abstinent days increased by half of a standard deviation from baseline ( $M = 0.56$ ) to the 9-month interview ( $M = 0.72; d = 0.51$ ), and DPDD declined more than half a standard deviation from baseline ( $d = 0.60$ ). At baseline participants reported, on average, using illicit drugs about 2 days per week ( $M = 0.28$ ) and by the 9-month interview this average value had declined to less than 1 day per week ( $M = 0.19; d = 0.27$ ).

The first aim of the study was to replicate Kelly et al.'s (2011) results, using our more externally valid sample. Table 3 displays the direct effects between 12-step attendance in months 0–3, S/R beliefs and practices in months 4–6, and substance use in months 7–9. Because of missing data and participants lost during follow-up, the number of participants reported in Tables 3 and 4 analyses range from 90 to 112. As shown, 12-step attendance significantly predicted an increase in days abstinent from alcohol ( $\beta = .24, p < .01$ ) and reduced DPDD ( $\beta = -.29, p < .01$ ), even after controlling for concurrent treatment for alcohol, drug, and emotional problems. No relationship was found, however, between frequency of 12-step attendance and later illicit drug use. The path between 12-step attendance and increased RBB scores was significant and positive ( $\beta = .19, p < .02$ ). Finally, positive changes in RBB scores were significantly predictive of both increased abstinent days ( $\beta = .25, p < .05$ ) and reduction in drinks per drinking day ( $\beta = -.25, p < .05$ ). However, there was no association between RBB scores and proportion of days participants took their most frequently used drug.

**Mediation**—Results showed that 15% of the direct effect of 12-step attendance on later frequency of abstinent days was predicted by participants' recent S/R beliefs and practices,  $Z = 1.83, p = .07$ . This finding is similar to results from Kelly et al.'s (2011) mediational analysis on their outpatient sample in which the RBB measure accounted for 14% of the effect of 12-step attendance on PDA in months 9–12. Regarding drinking intensity, recent S/R beliefs and practices accounted for 10% of the direct effect of 12-step attendance on DPDD from 6–9 months,  $Z = -1.70, p = .089$ .

Our second aim was to investigate the item content in the RBB measure that accounted for mediational effects. To begin, we computed the two RBB scales described by Connors et al. (1996), namely, Formal Religious Practices and God Consciousness, and examined the magnitude of the direct lagged pathways from 12-step meeting attendance in months 0–3 to each of these scale scores in the 6-month interview (months 4–6). We then analyzed the

paths from each of the RBB scales to substance use in months 7–9. As in our first analysis, we controlled for alcohol, drug, and psychological treatment received from baseline to 3 months. The association between 12-step meeting attendance and substance use was significant and is described above as part of our first mediation analysis. Table 4 displays the results of our second set of General Linear Modeling (GLM) analyses. Our first causal step analysis demonstrated that 12-step attendance in months 0–3 predicted increased Formal Religious Practices in months 4–6 ( $\beta = .13, p = .03$ ) but not increased God Consciousness ( $\beta = .14, p = .10$ ). Second, we found that Formal Religious Practices in months 4–6 did not predict either 9-month PDA ( $\beta = .17, p = .10$ ) or DPDD ( $\beta = -.17, p = .11$ ). However, increased God Consciousness significantly predicted both increased PDA ( $\beta = .29, p < .01$ ) and decreased DPDD ( $\beta = -.26, p = .02$ ). These results suggest that Formal Religious Practices increase as a result of increased 12-step attendance, but these do not predict later drinking outcomes. In contrast, God Consciousness is not influenced by 12-step attendance, but it does significantly predict improved drinking outcomes. The MGC scale in Table 4 shows our effort to resolve this situation. To modify the God Consciousness scale, we took one item, meditation practice, from the Formal Religious Practices scale and added it to the God Consciousness scale. In their original psychometric paper on the RBB, Connors et al. (1996) found that practicing meditation was not highly associated ( $B < 0.30$ ) with either the Formal Religious Practices or the God Consciousness factors. Their decision to add meditation practice to the Formal Religious Practices factor was based solely on their opinion regarding where this item would best fit. However, in our view, meditation practice seems similar to the content of the items in God Consciousness scale, such as prayer. Thus, we added meditation practice to the God Consciousness scale and reran our mediation analysis to evaluate whether this alteration would change the results.

Table 4 shows the results of shifting meditation practice to the God Consciousness scale. We found that this minor modification changed our findings, such that lagged 12-step attendance significantly predicted the modified indicator of God Consciousness ( $\beta = .20, p = .02$ ). The MGC also significantly predicted increased abstinence ( $\beta = .30, p < .01$ ) and decreased drinking intensity ( $\beta = -.09, p < .01$ ) at 9 months. Not presented in Table 4, secondary analyses were done to assess how the modified Formal Religious Practices scale (composed of two items after deleting meditation practice, namely, “holy scripture reading” and “formal worship attendance”) would perform in the causal step model. Our lagged analyses indicated that 12-step meeting attendance at 3 months was now unrelated to Formal Religious Practices at 6 months ( $p = .18$ ), and that Formal Religious Practices at 6 months did not predict increased PDA ( $p = .49$ ) or decreased DPDD ( $p = .26$ ) at 9 months. Consistent with our earlier analyses, none of the S/R measures predicted later illicit drug use.

**Mediation**—Analyses showed that MGC significantly mediated the effect of 12-step attendance on 9-month PDA ( $Z = 1.95, p < .05$ ), and that it accounted for 22% of the direct effect of AA attendance on later PDA. However, we found that MGC had only a marginally significant mediating effect on the association between 12-step attendance and reduced DPDD at 9 months, ( $Z = 1.90, p = .06$ ) and accounted for 14% of the variance of the direct effect of 12-step attendance on DPDD.

**Trajectory of Spiritual Change and 12-Step Attendance**—Hierarchical linear models (Raudenbush & Bryk, 2002) were used to assess the nature of change, at the individual level, in the MGC scores over 9 months and to assess whether the change over time in MGC scores was moderated by frequency of 12-step attendance. A total of 118 participants (91%) provided sufficient data to be included in these analyses. In our first analysis, we fit a level-1 random intercept model to estimate the amount of outcome variation that existed within versus between individuals. Results showed that the intercept was significantly different from zero, indicating that the degree of MGC of the average participant was nonzero;  $t(117) = 44.49, p < .001$ , MGC grand mean = 18.21, ( $SD = 5.14$ ). Next, we fit an unconditional growth model that included only intercept and time as level-1 random effects predictors of MGC scores at baseline, 3, 6, and 9 months. The time effect was not significant,  $t(117) = -0.101, p = .92$ , indicating that, on average, MGC scores did not change significantly over time. However, results showed that the rate of change in MGC varied significantly between individuals,  $\chi^2(117) = 173.63, p < .001$ . Thus, we fit a conditional growth model to investigate whether the frequency with which participants attended 12-Step meetings between intake and 3 months moderated the change in participants' MGC scores over time. The 12-step attendance moderator was computed by applying a median split to the 3-month 12-step attendance data. This resulted in a low 12-step attendance group that consisted of 56 participants and a high 12-step attendance group that included 62 participants. From baseline to the 3-month follow-up, the proportion of days on which the low attendance group went to 12-step meetings ranged from 0 to 0.06, which translates to a maximum of 5.4 days of meeting attendance out of 90 days ( $M = 0.02, SD = 0.02$ ). During this same period, the high attendance group went to 12-step meetings on 6% to 100% of days and, on average, attended one meeting about every other day ( $M = 0.45, SD = 0.31$ ). A comparison of the intercepts indicated that, on average, the low 12-step attendance group did not differ from the high 12-step attendance group on MGC scores at baseline,  $t(116) = 0.87, p = .39$ . However, a significant time by group interaction was observed such that the low attendance group reported modestly declining MGC scores over 9 months, ( $b = -.11, p = .10$ ), whereas members in the high attendance group reported statistically significant increases in MGC scores over 9 months,  $t(116) = 2.07, p = .04$ .

Our third aim was to investigate whether the association between MGC and improved drinking outcomes occurred simply because MGC sustains 12-step attendance. To test this hypothesis, we investigated whether MGC scores at 6 months predicted PDA and DPDD at 9 months. Two GLMs were conducted, one to evaluate PDA and one to evaluate DPDD. Several variables were controlled for including baseline values of the respective drinking measure, baseline MGC scores, gender, ethnicity, marital status, and employment status. The proportion of days of formal treatment reported at intake, 3, and 6-month interviews were also entered as covariates to isolate the effects of interest. Unlike prior analyses, we also included two additional covariates, namely, the proportion of days from 3 to 6 months and from 6 to 9 months on which participants attended 12-step meetings. Results indicated that after controlling for the proportion of days on which participants attended 12-step meetings from 6 to 9 months MGC at 6 months still significantly predicted increased 9-month PDA, ( $\chi^2(1) = 3.94, p < .05, \beta = .20$ ) and narrowly missed predicting 9-month DPDD ( $\chi^2(1) = 3.79, p = .052, \beta = .21$ ).

## DISCUSSION

Our results largely replicated Kelly et al.'s (2011) finding that reductions in drinking during early 12-step affiliation are explained, in part, by increased spiritual practices. Although the results of our mediational analyses differed in some key respects from Kelly and colleagues' results, the two fully lagged prospective sets of analyses yielded strikingly similar estimates of the magnitude of the associations between (1) 12-step attendance and later increases in spiritual beliefs and practices, and (2) increases in spirituality and later increases in abstinence and reduced drinking intensity. In addition, when we added spirituality as a mediator, the percent change in the association between 12-step attendance and later alcohol use was similar to the change that Kelly et al. found in their study. Specifically, in our study the association between 12-step attendance and proportion days abstinent lessened by 15%, and Kelly et al. found that this association lessened by 14%.

Our findings offer some important contributions to better understand the nature and function of spiritual practices in 12-step programs. First, our study used a sample with high external validity. Participants were ethnically diverse, had limited prior 12-step involvement, and had comorbid substance use and psychological disorders. They also had higher problem severity, less social support, and a higher unemployment rate than participants had in prior studies on spirituality and 12-step participation, e.g., Kelly et al. (2011). Given that alcoholics in the community are commonly faced with these problems, the findings from our study may provide a stronger demonstration that spiritual practices mediates 12-step benefit in real-world populations than do the findings of Kelly and colleagues. Second, unlike earlier investigations of the mediational effect of spirituality, our causal step models statistically controlled for formal treatment-seeking throughout study participation. Thus, our study may have more effectively isolated the effects of 12-step attendance and changes in spirituality on drinking outcomes.

Three different methods for defining spirituality using the RBB were examined in this study, and, collectively, our findings offer important insight into the functional role(s) of spirituality in 12-step programs. Applying Connors et al.'s (1996) two scale approach produced the weakest evidence for spiritual mediation. This occurred because only one item in the Formal Religious Practices scale was sensitive to the influence of AA meeting attendance, and because the sum of responses on this scale was not influential in predicting drinking behavior. In contrast, the items in the God Consciousness scale were predictive of drinking reduction but were not influenced by prior 12-step meeting attendance. Recent religious and spiritual practices, the score that Kelly et al. (2011) created by summing all recent formal worship practices and spiritual experiences reported in the RBB was sensitive enough to be influenced by the AA meeting attendance and served to predict drinking reduction. Our replication with a more externally valid sample confirmed that this composite measure of religious and spiritual practices partially mediated the association between 12-step participation and drinking outcomes. However, using this global score made it impossible to determine the relative value of formal religious practices, e.g., attending worship services, in explaining 12-step effects. Thus, we conducted a secondary analysis in which we compared the mediating effects of two RBB subscales, a Formal Religious Practices scale and an MGC scale. Results showed that frequency of 12-step attendance was

largely independent of frequency of later formal worship practices, even though participants reported frequent formal worship. Results also suggested that the practice of formal worship was relatively unimportant in predicting reduced drinking among early 12-step affiliates. However, our MGC scale score was significantly predicted by early 12-step attendance and served to predict both later proportion of days abstinent from alcohol and drinking intensity. In this light, Kelly et al.'s recent spiritual practices score that sums all recent formal religious practices and spiritual experiences, includes two items that are not influenced by the 12-Step experience and that do not appear to contribute to describing the functional role of spirituality in 12-step programs. For this reason, our MGC scale may be a more efficient indicator of spirituality that can be used in future 12-step research to investigate the functional roles of spirituality in 12-step programs.

Researchers do not often investigate the latency period between exposure to an active ingredient, mobilization of a change mechanism, and the positive action of a mediator on outcome. Our results indicated that the functional benefits of spiritual practices may occur quite early, between 4 and 6 months of 12-step affiliation, a finding that is consistent with findings in partially lagged studies (Robinson et al., 2007). Especially, intriguing was our finding that the initial intensity of 12-step attendance between baseline and 3 month defined the trajectory of later spiritual practices. Specifically, participants who initially attended 12-step meetings more frequently had sustained increases in MGC scores, whereas participants who infrequently attended 12-step meetings had a steady decline in spiritual beliefs and practices over 9 months. This finding raises the possibility that participants' initial motivation and willingness to attend 12-step meetings is an important prerequisite that influences the mobilization, timing, and functional benefit of spirituality as it is practiced in 12-step programs.

No support was found for the hypothesis that increased spiritual practices improved drinking outcomes because such changes enhanced 12-step meeting attendance. Specifically, we found that changes in our MGC measure continued to account for later increased abstinence even after controlling for prior and concurrent 12-step attendance. The same pattern was observed for the effects of spiritual change on reduced drinking intensity, albeit at the trend level ( $p < .052$ ). Although the possibility remains that spiritual practices mobilize other prescribed 12-step activities that actually explain drinking reduction, we consider this possibility to be remote. Especially, during early 12-step affiliation, for example, correlations are positive and strong between 12-step meeting attendance and composite measures of commitment to, and involvement in, prescribed 12-step practices (e.g., Tonigan, Connors, & Miller, 2003). As such, statistical control for 12-step meeting attendance provided a strong signal about the likelihood that other measures of 12-step engagement would be influenced by reported changes in spirituality.

Our findings are perplexing about the illicit drug use of early AA affiliates. On one hand, the percentage of participants who abstained from illicit drugs increased from 30% at baseline to the 53% at the 9-month interview, and our continuous measure of illicit drug use indicated a modest pre-post reduction in the frequency of drug use ( $d = 0.27$ ). This finding is consistent with the message of abstinence from all mood altering drugs that is endorsed by all 12-step programs. Contrary to many reports, however, this reduction in illicit drug use was not



related to frequency of AA meetings (e.g., Gossop et al., 2007; Timko et al., 2006; Timko & Sempel, 2004; Weiss et al., 2005; Witbrodt & Kaskutas, 2005; Worley et al., 2008). Likewise, changes in spirituality were not predictive of later illicit drug use, although most of the spirituality measures did predict increased alcohol abstinence and reduced drinking intensity. Whether our findings are an artifact of how we measured illicit drug use, self-selected characteristics of our sample, or the timing of assessments is unclear but each should be acknowledged as a possibility. We encourage future investigations of spiritual change among 12-step affiliates to include measures of illicit drug use to further assess the specificity of the spirituality meditational effect.

### Study's Limitations

Several study limitations should be noted. First, we intentionally recruited adults with limited AA histories to restrict past learning effects associated with AA exposure. While enhancing the internal validity of the study, this decision also limits the generalizability of our findings. In particular, a majority of adults seeking substance use treatment have extensive AA histories and it is unclear if these cumulative experiences may alter the nature of findings in this study. Second, we used frequency of self-reported AA meeting attendance as our indicator of AA exposure. Our analyses indicated that this measure was reliable, but attendance and commitment to AA prescribed beliefs and practices are quite different. Equally troubling, evidence suggests that AA meetings differ in perceived social dynamics, and that these differences predict increased abstinence (Rice & Tonigan, 2012). Collectively, future research should investigate the linkages of interest in this study with measures that are more sensitive to both personal commitment and practice of AA prescribed activities and the nature of AA meetings.

### CONCLUSIONS

In sum, our study offers the strongest evidence to date that changes in spiritual practices are associated with an increased abstinence from alcohol consumption among 12-step affiliates. Our sample had high external validity, and findings did not support the prediction that spiritual practices indirectly accounted for increased abstinence by simply sustaining 12-step meeting attendance. Why changes in spiritual practices were associated with increased abstinence remains unclear, however. Our efforts to better define spiritual practices offer some clues in this matter. Specifically, we found that 12-step attendance did predict increased formal religious practices like attending worship services, but that such increases did not predict increased abstinence nor did they explain 12-step related benefit. Changes in the practice of prayer and meditation, however, did account for increased abstinence and reduced drinking intensity. Strong parallels can be identified between these findings and recent work on the active ingredients and mediators of mindfulness-based therapy (e.g., Witkiewitz, Marlatt, & Walker, 2005), which posits that meditation may produce enhanced cognitive and behavioral coping skills that replace unhealthy with healthy responses to negative emotion and drinking cues such as nonjudgmental and present-centered awareness. We believe that high priority should be assigned to investigating the causal linkages between changes in spiritual practices, cognitive and behavioral coping skills, negative emotional

arousal as more specific mechanisms by which involvement in 12-step programs contributes to ultimate drinking outcomes.

## Acknowledgments

This research was supported by grants K02-AA00326 and R21AA016974 from the National Institute on Alcohol Abuse and Alcoholism.

## Biographies



**J. Scott Tonigan**, PhD, is a Research Professor at the Department of Psychology, University of New Mexico, Albuquerque, NM, USA, and at the Center on Alcoholism, Substance Abuse, and Addictions (CASAA), University of New Mexico. His research interests include community-based mutual-help programs, processes of change in addiction treatment, and the measurement of substance use and consequences. Dr. Tonigan has published more than 150 papers published in addiction research and he has 20-year career as a funded researcher for the National Institute on Alcohol Abuse and Alcoholism (NIAAA).



**Kristina N. Rynes**, PhD, is a postdoctoral fellow at the Center on Alcoholism, Substance Abuse, and Addictions, University of New Mexico, Albuquerque, NM, USA. Her research interests include community-based mutual-help programs, processes of change in treatment, and the role that family and social systems play in the maintenance and change of alcohol and substance misuse. She has written articles on mediators and moderators of change in 12-step programs, for instance, “Do Social Networks Explain 12-Step Sponsorship Effects? A Prospective Lagged Mediation Analysis” with J. Scott Tonigan in *Psychology of Addictive Behaviors* (2011, September 5).



**Barbara S. McCrady**, PhD, is a Distinguished Professor of Psychology and the Director of the Center on Alcoholism, Substance Abuse, and Addictions, University of New Mexico, Albuquerque, NM, USA. Her research interests include the development and testing of efficacious treatments for persons with substance use disorders (SUDs), treatments for

women with SUDs, and the role of family and other social support systems in treatment. Dr. McCrady has published more than 225 articles, chapters, and books in the addictions field. Her research has been funded by the National Institutes of Health since 1979.

## GLOSSARY

<b>Spiritual awakening</b>	Within AA spiritual awakening refers to the moment that a member recognizes his/her connection or relationship with a higher power. Although the 12th step of AA explicitly states that such an awakening occurs as a result of working the preceding 11 steps, many members report having a spiritual awakening quite early in their AA affiliation
<b>Spiritual growth</b>	Within AA spiritual growth refers to the belief that the deepening of a relationship with a higher power is a lifelong process and should be assigned a high priority. Central to this construct is the belief that spiritual stagnation places an individual in jeopardy of an alcohol relapse
<b>Spiritual transformation</b>	Within AA spiritual transformation refers to a fundamental shift in character or personality that occurs because of a deepening relationship with a higher power. Specifically, the core AA literature describes the alcoholic as egocentric and selfish to the extreme, traits that precede the onset of alcohol dependence but which become further magnified as alcohol dependency becomes established. Through spiritual awakening and growth, these underlying personality traits are thought to be replaced with selflessness and consideration of others.

## REFERENCES

- Alcoholics Anonymous World Services. Alcoholics Anonymous: Big book. 3rd. New York, NY: Author; 2001.
- American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th. Washington, DC: Author; 2000. Text Revision
- Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*. 1986; 51(6):1173–1182. [PubMed: 3806354]
- Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*. 1994; 6(4):284–290.
- Connors G, Tonigan JS, Miller WR. The religious background and behavior instrument: Psychometric and normed findings. *Psychology of Addictive Behaviors*. 1996; 10:90–96.
- Emrick, CD.; Tonigan, JS.; Montgomery, HA.; Little, L. Alcoholics Anonymous: What is currently known?. In: McCrady, BS.; Miller, WR., editors. *Research on Alcoholics Anonymous: Opportunities and alternatives*. Piscataway, NJ: Rutgers Center for Alcohol Studies; 1993. p. 41-76.
- Gossop M, Stewart D, Marsden J. Readiness for change and drug use outcomes after treatment. *Addiction*. 2007; 102(2):301–308. [PubMed: 17222285]
- Hunter, JE.; Schmidt, FL. *Methods of meta-analysis: Correcting error and bias in research findings*. 2nd. Thousand Oaks, CA: Sage; 2004.
- Kaskutas LA, Bond J, Humphreys K. Social networks as mediators of the effect of Alcoholics Anonymous. *Addiction*. 2002; 97(7):891–900. [PubMed: 12133128]

- Kaskutas LA, Turk N, Bond J, Weisner C. The role of religion, spirituality and Alcoholics Anonymous in sustained sobriety. *Alcoholism Treatment Quarterly*. 2003; 21(1):1–16.
- Kelly J, Moos R. Dropout from 12-step self-help groups: Prevalence, predictors, and counteracting treatment influences. *Journal of Substance Abuse Treatment*. 2003; 24(3):241–250. [PubMed: 12810145]
- Kelly JF, Magill M, Stout RL. How do people recover from alcohol dependence? A systematic review of the research on mechanisms of behavior change in Alcoholics Anonymous. *Addiction Research & Theory*. 2009; 17(3):236–259.
- Kelly JF, Stout RL, Magill M, Tonigan JS, Pagano ME. Spirituality in recovery: A lagged mediational analysis of Alcoholics Anonymous' principal theoretical mechanism of behavior change. *Alcoholism: Clinical and Experimental Research*. 2011; 35(3):454–463.
- Kelly JF, Stout RL, Tonigan JS, Magill M, Pagano ME. Negative affect, relapse, and Alcoholics Anonymous (AA): Does AA work by reducing anger? *Journal of Studies on Alcohol and Drugs*. 2010; 71(3):434–444. [PubMed: 20409438]
- Kelly JF, Yeterian JD, Myers MG. Treatment staff referrals, participation expectations, and perceived benefits and barriers to adolescent involvement in twelve-step groups. *Alcoholism Treatment Quarterly*. 2008; 26(4):427–449.
- Laffaye C, McKeller JD, Ilgen MA, Moos RH. Predictors of 4-year outcome of community residential treatment for patients with substance use disorders. *Addiction*. 2008; 103:671–680. [PubMed: 18339113]
- MacKinnon DP, Lockwood CM, Hoffman JM, West SG, Sheets V. A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods*. 2002; 7(1):83–104. [PubMed: 11928892]
- Miller, WR. NIAAA Project MATCH Monograph Series. Vol. 5. Rockville, MD: National Institute on Alcohol Abuse and Alcoholism; 1996. Form 90: A structured assessment interview for drinking and related behaviors.
- Morgenstern J, Labouvie E, McCrady BS, Kahler CW, Frey RM. Affiliation with Alcoholics Anonymous after treatment: A study of its therapeutic effects and mechanisms of action. *Journal of Consulting and Clinical Psychology*. 1997; 65(5):768–777. [PubMed: 9337496]
- Oakes KE. Purpose in life: A mediating variable between involvement in Alcoholics Anonymous and long-term recovery. *Alcoholism Treatment Quarterly*. 2008; 26(4):450–463.
- Pagano M, Friend K, Tonigan J, Stout R. Helping other alcoholics in Alcoholics Anonymous and drinking outcomes: Findings from project MATCH. *Journal of Studies on Alcohol*. 2004; 65(6): 766–773. [PubMed: 15700515]
- Preacher KJ, Hayes AF. SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers*. 2004; 36(4):717–731.
- Project MATCH Research Group. Matching alcoholism treatments to client heterogeneity: Project MATCH posttreatment drinking outcomes. *Journal of Studies on Alcohol*. 1997; 58(1):7–29. [PubMed: 8979210]
- Project MATCH Research Group. Matching alcoholism treatment to client heterogeneity: Treatment main effects and matching effects on within-treatment drinking. *Journal of Studies on Alcohol*. 1998; 59:631–639. [PubMed: 9811084]
- Raudenbush, SW.; Bryk, AS. Hierarchical linear models: Applications and data analysis methods. 2nd. Newbury Park, CA: Sage; 2002.
- Rice SL, Tonigan JS. Impressions of Alcoholics Anonymous (AA) group cohesion: A case for a non-specific factor predicting later AA attendance. *Alcoholism Treatment Quarterly*. 2012; 30(1):40–51. [PubMed: 25089071]
- Robinson E, Cranford J, Webb J, Brower K. Six-month changes in spirituality, religiousness and heavy drinking in a treatment-seeking sample. *Journal of Studies on Alcohol and Drugs*. 2007; 68(2): 282–290. [PubMed: 17286347]
- Ross HE. Benzodiazepine use and anxiolytic abuse and dependence in treated alcoholics. *Addiction*. 1993; 88(2):209–218. [PubMed: 8106063]
- Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*. 1979; 86(2):420–428. [PubMed: 18839484]

- Terry, DJ.; Hog, MA.; White, KM. Attitude-behavior relations: Social identity and group membership. In: Terry, DJ.; Hogg, MA., editors. Attitudes, behavior, and social context: The role of norms and group membership. Mahwah, NJ: Lawrence Erlbaum Associates; 2000. p. 67-94.
- Timko C, Billow R, DeBenedetti A. Determinants of 12-step group affiliation and moderators of the affiliation-abstinence relationship. *Drug and Alcohol Dependence*. 2006; 83(2):111–121. [PubMed: 16338102]
- Timko C, Finney JW, Moos RH. The 8-year course of alcohol abuse: Gender differences in social context and coping. *Alcoholism: Clinical and Experimental Research*. 2005; 29:612–621.
- Timko C, Sempel JM. Intensity of acute services, self-help attendance and one-year outcomes among dual diagnosis patients. *Journal of Studies on Alcohol*. 2004; 65(2):274–282. [PubMed: 15151360]
- Tonigan JS. Changing AA practices and outcomes: Project MATCH three-year follow-up. *Alcoholism: Clinical and Experimental Research*. 2003; 27(3):526–528.
- Tonigan JS, Beatty G. Twelve-step program attendance and polysubstance use: Interplay of alcohol and illicit drug use. *Journal of Studies on Alcohol and Drugs*. 2011; 72(5):864–871. [PubMed: 21906513]
- Tonigan, JS.; Connors, GJ.; Miller, WR. Participation and involvement in Alcoholics Anonymous. In: Del Boca, FK., editor. Treatment matching in alcoholism. New York, NY: Cambridge University Press; 2003. p. 184-204.
- Tonigan JS, Rice SL. Is it beneficial to have an Alcoholics Anonymous sponsor? *Psychology of Addictive Behaviors*. 2010; 24(3):397–403. [PubMed: 20853924]
- Tonigan JS, Miller WR, Brown JM. The reliability of form 90: An instrument for assessing alcohol treatment outcome. *Journal of Studies on Alcohol*. 1997; 58(4):358–364. [PubMed: 9203116]
- Tonigan JS, Toscova R, Miller WR. Meta-analysis of the literature on alcoholics anonymous: Sample and study characteristics moderate findings. *Journal of Studies on Alcohol*. 1996; 57(1):65–72. [PubMed: 8747503]
- Tsuang JW, Shapiro E, Smith TL, Schuckit MA. Drug use among primary alcoholic veterans. *The American Journal of Drug and Alcohol Abuse*. 1994; 20(4):483–493. [PubMed: 7832181]
- Weiss RD, Griffin ML, Gallop RJ, Najavits LM, Frank A, Crits-Christoph P, et al. The effect of 12-step self-help group attendance and participation on drug use outcomes among cocaine-dependent patients. *Drug and Alcohol Dependence*. 2005; 77(2):177–184. [PubMed: 15664719]
- Westerberg VS, Miller WR, Tonigan JS. Reliability of form 90D: An instrument for quantifying drug use. *Substance Abuse*. 1998; 19(4):179–189. [PubMed: 12511815]
- Witbrodt J, Kaskutas LA. Does diagnosis matter? Differential effects of 12-step participation and social networks on abstinence. *American Journal of Drug and Alcohol Abuse*. 2005; 31(4):685–707. [PubMed: 16320441]
- Witkiewitz K, Marlatt GA, Walker D. Mindfulness-based relapse prevention for alcohol and substance use disorders. *Journal of Cognitive Psychotherapy: An International Quarterly*. 2005; 19(3):211–225.
- Worley M, Gallop R, Gibbons MBC, Ring-Kurtz S, Present J, Weiss RD, et al. Additional treatment services in a cocaine treatment study: Level of services obtained and impact on outcome. *American Journal on Addictions*. 2008; 17(3):209–217. [PubMed: 18463998]
- Zemore SE. A role for spiritual change in the benefits of 12-step involvement. *Alcoholism: Clinical and Experimental Research*. 2007; 31(3):76–79.

TABLE 1

Meta-analysis of spiritual practices as a change mechanism in 12-step programs

Study	Design	Causal path model			N <sup>a</sup>
		AA to abstinence	AA to S/R	S/R to abstinence	
Kaskutas et al. (2003) <sup>b</sup>	Partial lagged	$B = 0.27$	—	$B = 0.25$	587
Kelly et al. (2011)	Fully lagged	$B = 0.24, A^c$	$B = 0.25, A$	$B = 0.18, A$	774
	Fully lagged	$B = 0.21, O^c$	$B = 0.14, O$	$B = 0.21, O$	952
Oakes (2008) <sup>d</sup>	Cross-sectional	$B = 0.41$	$B = 0.23$	$B = 0.21$	77
Robinson et al. (2007) <sup>b</sup>	Partial lagged	$B = 0.02$	—	$B = 0.05$	123
Tonigan (2003)	Partial lagged	$B = 0.27, O$	$B = 0.25, O$	$B = -0.06, O$	637
Zemore (2007) <sup>b</sup>	Partial lagged	$B = 0.43$	$B = 0.29$	$B = 0.10$	537
$B_{\text{weighted}} (SD)$		0.27 (0.04)	0.22 (0.06)	0.14 (0.04)	526.71
Chi-square homogeneity		$p < .001$	$p < .03$	$p < .001$	
% sampling error <sup>e</sup>		22.37%	45.02%	16.51%	

Note: AA = Alcoholics Anonymous. S/R = Recent spiritual and religious beliefs and practices.

<sup>a</sup> Number of participants that contributed to the estimation of path coefficients.

<sup>b</sup> Reported odds ratios were converted to unstandardized regression coefficients by multiplying them by the natural log. The resulting coefficients were then standardized via  $\beta = b (s_x/s_y)$ . Adjusted (AOR) and unadjusted odd ratios (OR) were computed using identical procedures, and therefore, the resultant Beta weights do not reflect covariate adjustments made in AOR. In cases where standard deviations were not reported by authors (or could not be calculated from data presented in the report) the values that were provided in the psychometric papers were substituted.

<sup>c</sup> Project MATCH Research Group (1998) aftercare sample coded A, and Project MATCH outpatient sample coded O.

<sup>d</sup> Author reported bivariate correlations which are equivalent to Beta.

<sup>e</sup> Percentage of sampling error represents the proportion of variance in the aggregated betas that is not due to systematic “third” variables.



**TABLE 2**Study participant help seeking and substance use: Baseline and follow-up ( $N = 130$ )

	<b>Intake</b>	<b>3</b>	<b>6</b>	<b>9</b>
Help seeking				
Proportion days 12-step attendance, $M (SD)$	0.17 (0.20)	0.24 (0.31)	0.23 (0.29)	0.17 (0.24)
% Attending 12-step	100	79.7	78.9	76.1
Proportion days treatment, $M (SD)$	0.07 (0.14)	0.12 (0.19)	0.11 (0.20)	0.09 (0.17)
% Attending treatment	58.9	64.4	58.8	55.8
Days religious attendance, $M (SD)$	3.76 (8.00)	6.14 (15.03)	4.57 (8.02)	3.88 (9.97)
Days formal education, $M (SD)$	3.84 (13.05)	3.20 (10.12)	3.32 (10.07)	5.60 (21.25)
Substance use				
PDA (alcohol), $M (SD)$	0.56 (0.31)	0.72 (0.32)	0.74 (0.30)	0.72 (0.32)
% Abstinent from alcohol	0	19.5	24.6	25.7
Drinks per drinking day (DPDD), $M (SD)$	14.01 (9.13)	9.67 (10.78)	8.84 (9.05)	8.60 (8.80)
Proportion days most frequently used illicit drug	0.28 (0.33)	0.22 (0.32)	0.22 (0.33)	0.19 (0.30)
% Abstinent from illicit drugs	30.2	49.2	47.4	53.1

**TABLE 3**

Summary of fully lagged causal step General Linear Modeling (GLM) analyses: RBB recent spiritual practices and beliefs

	<i>B</i> ( <i>SE</i> )	$\beta$	$\chi^2$	<i>p</i>
Path: AA predicting Substance use				
PDA (alcohol)	0.41 (0.15)	.24	7.14	.01
DPDD	-1.63 (0.52)	-.29	9.84	.00
Proportion most frequently used illicit drug	0.06 (0.08)	.06	0.23	.48
Path: AA predicting RBB				
	6.39 (2.70)	.19	5.60	.02
Path: RBB predicting substance use				
PDA (alcohol)	0.0139 (0.01)	.25	5.80	.02
DPDD	-0.04 (0.02)	-.25	4.92	.03
Proportion days most frequently used illicit drug	0.00 (0.00)	.03	0.04	.84

*Note:* RBB = Religious Background and Behavior scale (Connors et al., 1996). *B* = unstandardized coefficient,  $\beta$  = standardized coefficient. AA = Alcoholics Anonymous. PDA = Proportion days abstinent. DPDD = Drinks per drinking day. The following covariates were included in all analyses: gender, age, ethnicity, marital status, employment status, baseline value of the dependent measure, and formal treatment for substance use or psychological problems. All formal treatment received prior to the study as well as concurrent treatment experiences were included as control variables.

TABLE 4

Summary of fully lagged causal step GLM analyses: RBB scales

	<i>B</i> ( <i>SE</i> )	$\beta$	$\chi^2$	<i>p</i>
Path: AA predicting RBB scales				
Formal religious practices	4.55 (2.07)	.13	4.85	.028
God consciousness	1.84 (1.12)	.14	2.70	.101
Modified god consciousness	3.53 (1.45)	.20	5.93	.015
Path: RBB scales predicting substance use				
Formal religious practices				
PDA (alcohol)	0.01 (.01)	.17	2.64	.10
DPDD	-0.04 (.02)	-.17	2.54	.11
Proportion days use of most frequently used drug	0.00 (0.01)	.03	0.04	.84
God consciousness				
PDA (alcohol)	0.04 (0.01)	.29	8.30	.00
DPDD	-0.11 (0.05)	-.26	5.72	.02
Proportion days use of most frequently used drug	0.00 (0.01)	.01	0.03	.87
Modified God consciousness				
PDA (alcohol)	0.03 (0.01)	.30	8.86	.00
DPDD	-0.09 (0.03)	-.29	7.78	.01
Proportion days use of most frequently used drug	0.00 (0.01)	.02	0.03	.86

*Note:* RBB = Religious Background and Behavior scale (Connors et al., 1996). *B* = unstandardized coefficient,  $\beta$  = standardized coefficient. PDA = Proportion days abstinent. DPDD = Drinks per drinking day. The following covariates were included in all analyses: gender, age, ethnicity, marital status, employment status, baseline value of the dependent measure, and formal treatment for substance use or psychological problems. All formal treatment received prior to the study as well as concurrent treatment experiences were included as control variables.