Published in final edited form as:

Psychol Addict Behav. 2013 September ; 27(3): 553–561. doi:10.1037/a0029268.

The General Alcoholics Anonymous Tools of Recovery: The Adoption of 12-Step Practices and Beliefs

Brenna L. Greenfield and J. Scott Tonigan

Center on Alcoholism, Substance Abuse, and Addictions (CASAA) and Department of Psychology, University of New Mexico

Abstract

Working the 12 steps is widely prescribed for Alcoholics Anonymous (AA) members although the relative merits of different methods for measuring step-work have received minimal attention and even less is known about how step-work predicts later substance use. The current study (1) compared endorsements of step-work on an face-valid or direct measure, the Alcoholics Anonymous Inventory (AAI), with an indirect measure of step-work, the General Alcoholics Anonymous Tools of Recovery (GAATOR), (2) evaluated the underlying factor structure of the GAATOR and changes in step-work over time, (3) examined changes in the endorsement of step-work over time, and (4) investigated how, if at all, 12-step-work predicted later substance use. New AA affiliates ($N = 130$) completed assessments at intake, 3, 6, and 9 months. Significantly more participants endorsed step-work on the GAATOR than on the AAI for nine of the 12 steps. An exploratory factor analysis revealed a two-factor structure for the GAATOR comprising Behavioral Step-Work and Spiritual Step-Work. Behavioral Step-Work did not change over time, but was predicted by having a sponsor, while Spiritual Step-Work decreased over time and increases were predicted by attending 12-step meetings or treatment. Behavioral Step-Work did not prospectively predict substance use. In contrast, Spiritual Step-Work predicted percent days abstinent, an effect that is consistent with recent work on the mediating effects of spiritual growth, AA, and increased abstinence. Behavioral and Spiritual Step-Work appear to be conceptually distinct components of step-work that have distinct predictors and unique impacts on outcomes.

Keywords

Alcoholics Anonymous; self help; mutual help; 12 Steps; substance use

Alcoholics Anonymous (AA) is the largest community-based mutual-help program for alcohol-related problems, and 12-Step Treatment, which is based upon AA principles and practices, is one of the most popular therapeutic approaches in the United States (Substance Abuse and Mental Health Services Administration, 2008). Both 12-step and non 12-step treatment providers routinely refer patients to AA during and after treatment (Roman & Johnson, 2004), and evidence suggests that AA participation is associated with increased abstinence (e.g., Moos & Moos, 2006; Tonigan & Rice, 2010). Understandably, then, investigations have sought to understand *why* AA is beneficial. In general, change processes that explain AA-related benefit include increases in abstinence self-efficacy (for a review, Forcehimes & Tonigan, 2008), social support for abstinence (for a review, see Groh, Jason, & Keys, 2008), and spirituality (e.g., Kaskutas, Turk, Bond, & Weisner, 2003; Kelly, Stout,

Magill, Tonigan, & Pagano, 2011; Robinson, Cranford, Webb, & Brower, 2007; Zemore, 2007).

In spite of the advances in understanding why AA may be helpful, we still know very little about the prescribed 12-step practices that mobilize beneficial change processes. Of the literally hundreds of empirical studies that have been conducted on AA, for example, only a few evidence-based recommendations can be made regarding what new AA affiliates ought to do to enhance the likelihood of increased abstinence. Strong support, for example, has been found for the benefits of acquiring a 12-step sponsor, especially during early affiliation (e.g., Bond, Kaskutas, & Weisner, 2003; Tonigan & Rice, 2010; Witbrodt & Kaskutas, 2005). Second, even after controlling for self-selective factors like motivation, other prescribed AA activities, and concurrent formal treatment (e.g., Kelly et al., 2010; Moos & Moos, 2006; Tonigan & Beatty, 2011) evidence is clear that *sustained* AA meeting attendance is predictive of increased abstinence. And, finally, several investigative teams have reported that service work, sometimes broadly defined, is a strong predictor of increased alcohol abstinence among 12-step affiliates (Pagano et al., 2010; Zemore, Kaskutas, & Ammon, 2004).

Step-work, the process of systematically working each of the 12 steps, is a cornerstone of the AA program, and it is widely encouraged during 12-step treatment. At the same time, empirical evidence for its utility is sparse (e.g., Carroll, 1993; Gilbert, 1991) and efforts to document its salubrious properties have been hampered by conceptual and measurement issues. Two approaches have been applied to assess the utility of working the 12 steps. On one hand, in addition to assessing an array of 12-step related behaviors and beliefs, multidimensional measures frequently ask respondents directly whether they have *completed* specific steps (yes/no; e.g., Brown & Peterson, 1991; Humphreys, Kaskutas, & Weisner, 1998; Kelly, Urbanowski, Hoepfner, & Slaymaker, 2011; Tonigan, Connors, & Miller, 1996). While facevalid, this approach relies upon a subjective evaluation about what constitutes the “completion” of a step, and there is reason to suspect that the criteria used to make this evaluation are not consistent across individuals. Also problematic, when “completion” of a step is reported, this method assumes that all beliefs and practices embodied in a given step are uniformly endorsed and internalized. In spite of these concerns, this is an efficient and popular method to collect information and its lack of precision is often considered to be minimized when such information is aggregated into larger composite measures indicating commitment to, and practice of, other prescribed 12-step behaviors (Connors, Tonigan, & Miller, 2002; Johnson, Finney, & Moos, 2006; Krentzman et al., 2011; Ouimette, Finney, Gima, & Moos, 1999; Owen et al., 2003).

Gilbert (1991) introduced a second and indirect method to quantify step work among 12-step affiliates, a method that circumvented the need for subjective evaluations of step completion. Specifically, Gilbert developed 42 statements about the beliefs and practices included in the first three steps, and the extent of respondent endorsement of these step-specific statements was used to infer the adoption and internalization of these steps. As an example, “I cannot control my use of alcohol” is an item from the Step 1 *Powerlessness* scale. Carroll (1993) applied this approach to the measurement of practicing Steps 11 and 12, and extended it to include respondent affective reactions to practicing these two steps, e.g., I feel “peace of mind.” Readily apparent, indirect measures of step-work have the distinct advantage that they allow for the partial endorsement of beliefs and/or practices nested within a step. Also important, indirect measures do not require higher-order cognitive appraisals about whether a step has been “completed.” The indirect approach, however, has limitations. Foremost, the utility of working *specific* steps is difficult to evaluate with indirect measures, with the net effect that this approach cannot unambiguously guide the selection of specific steps to be encouraged in treatment. Second, many beliefs and practices

enumerated in the 12 steps are routinely discussed in 12-step meetings. As such, it can be problematic to determine if the endorsement of a given practice or belief is the result of formally working the step per se or, alternatively, the result of attending 12-step discussion meetings. At an even broader level, some practices and beliefs abstracted from the 12-steps are not unique to 12-step programs and can be readily endorsed by individuals having no previous exposure to AA, e.g., sharing of a spiritual experience.

Recognition of “false positives” associated with indirect measures raises a fundamental issue about why step-work is even measured. On one hand, formal working of the steps is strongly encouraged in 12-step literature and in meetings and may indicate an important commitment by the 12-step member to engage in 12-step prescribed activities. Declared completion of steps (direct method) would be the least ambiguous method to make this determination. In studying 12-step programs, however, the measurement of step-work may be important only to the extent that prescribed practices and beliefs are internalized, guide actions, and predict drinking outcome. Here, the indirect method has clear advantages and the concern about false positives is substantially reduced.

The General Alcoholics Anonymous Tools of Recovery scale (GAATOR; Montgomery, Miller, & Tonigan, 1995) was developed as an indirect measure of the endorsement and internalization of the practices and beliefs enumerated in all of the 12 steps. The GAATOR has 24 Likert scale items, all of which are on a four-point scale. Each item maps onto a particular step (e.g., Step 12 - “I have told others about my spiritual experience”), but items are not explicitly labeled as a particular step. Abstracting practices and beliefs from the 12-steps was done in a manner that minimized author interpretation of latent principles (see Appendix). Frequently, for example, GAATOR items simply re-state a discrete idea presented within a step. Initial efforts to construct GAATOR items by the measure’s senior author, Dr. Montgomery, were first reviewed and refined by his co-authors Drs. Miller and Tonigan. Items were then reviewed by several AA members to determine the fit between GAATOR items and the step they were intended to represent.

To date, one preliminary psychometric analysis has investigated the scale structure of the GAATOR (Tonigan, Miller & Vick, 2000). These efforts were conducted as secondary analyses of data collected in the context of randomized clinical trials where 12-step attendance was low and engagement in other 12-step prescribed behaviors was reported by only a minority of the participants. Using Principal Components Analysis (PCA) the authors reported that the GAATOR consisted of three scales labeled *Higher Power*, *Self-Inventory*, and *Other Inventory*. It seems plausible that the GAATOR scale structure may be quite different for more 12-step involved problem drinkers who have not been enrolled in an NIH funded clinical trial.

This study had four aims. First, we sought to determine the extent of agreement between a face-valid measure that directly asked about step completion and an indirect measure of step-work, the GAATOR. High agreement between the two methods would argue against the development and/or use of indirect measures because of the additional burden required to complete a longer assessment. Our second aim was to conduct an exploratory factor analysis of an indirect measure of step work, the GAATOR, using a sample of adults who were new to AA, actively attending 12-step meetings, and were not participating in a larger NIH clinical trial. Third, we also investigated how, if at all, the structural endorsement of step work practices and beliefs changed over time. This aim was largely exploratory in nature but was included to provide a unique longitudinal perspective on step-work. And, fourth, we investigated whether self-reported step-work practices and beliefs prospectively predicted substance use outcomes, with substance use defined as drinks per drinking day, percentage of alcohol abstinent days, percentage of illicit drug use days, and a categorical

measure of alcohol abstinence. Aside from assessing the underlying structure of the GAATOR items (Aim 2), to our knowledge the study aims had not been addressed in prior research.

Method

Sample

This single-group naturalistic study investigated early affiliation with 12-step programs with special attention directed to the mechanisms accounting for 12-step related benefit (R21AA016974-01, Tonigan). Determined prior to consent by a research assistant, study eligibility criteria included (1) current DSM-IV alcohol abuse or dependence, (2) alcohol use in the past three months, and (3) attending at least one AA meeting in the past three months. Individuals were not eligible if they reported more than 16 weeks of lifetime AA meeting exposure or one year or more of abstinence from alcohol post-alcohol use disorder onset. Presence of a drug use disorder was not an exclusion criterion.

Participants were recruited in a large Southwestern city. Recruitment sources included substance use disorder treatment programs, community AA meetings, and public fliers/referrals from other participants. The University of New Mexico Institutional Review Board approved the study (#27147), and all participants provided written informed consent.

Procedure

No intervention was provided in this assessment-only study. However, participants were given referrals to treatment programs if interested or if study staff deemed it to be appropriate. Participants received \$50 for each completed assessment, for a maximum total of \$200. Participants were given a Breathalyzer test to verify that their BAC was .05 or less prior to each assessment. Participants with a BAC above .05 were asked to stay in the research offices until their BAC dropped to .05 or below, at which point they could either continue the research assessment or reschedule the appointment. To support study retention, participants provided contact information for one locator at intake whom research staff could contact in the event that participants could not be contacted to schedule a follow-up interview. This information was updated at each follow-up interview, as needed. Participants were also given cab or bus fare to travel home. Following the initial baseline assessment, follow-up assessments occurred at 3, 6, and 9 months. Three-month data were collected for 118 (91%) of the participants, with 113 of these being done in-person (96%). Six- and nine-month follow-up rates were 88% ($n = 114$, 108 in-person, 95%) and 88% ($n = 113$, 106 in-person, 94%), respectively. Attrition analyses indicated that baseline measures of alcohol use and dependency were not predictive of follow-up status (interviewed/lost) at any of the three follow-up interviews.

Measures

General Alcoholics Anonymous Tools of Recovery—The General Alcoholics Anonymous Tools of Recovery (GAATOR; Montgomery, Miller, & Tonigan, 1995) is a 24-item self-report measure of behaviors and practices associated with the 12 steps of AA. Each GAATOR item corresponds to a particular step (e.g., “I have turned my will and life over to my Higher Power” represents Step 3; “I have prayed and meditated” represents Step 11; see Appendix for further items), and for each step, a range of one to three GAATOR items are summed to represent “working a step” (see Table 1). Four response choices indicate the occurrence of each item in the past 90 days (definitely false, false, true, and definitely true).

A conservative estimate of completed GAATOR step-work was used for the comparison between the GAATOR and the Alcoholics Anonymous Inventory (AAI). For each item,

GAATOR responses of definitely false and false were recoded as 0, and responses of true and definitely true were recoded as 1. All items included in a step had to be endorsed as true or definitely true (i.e., “1”) for a step to be considered “completed.” Ratings of internal consistency for each step ranged from 0.42–0.85 (see Table 1). For all other analyses, a continuous measure of completed step-work was used. Item responses were recoded as 0 (definitely false or false), 1 (true), or 2 (definitely true). All items included in each step were summed for a total score ranging from 0 to 2 for steps with one GAATOR item (Step 3), 0 to 4 for steps with two items (Steps 1, 4–12), and 0 to 6 for steps with three GAATOR items (Step 2).

Alcoholics Anonymous Inventory—The Alcoholics Anonymous Inventory (AAI; Tonigan, Connors, & Miller, 1996) is a 13-item self-report instrument that measures recent and lifetime AA participation. In a sample of treatment seekers, it had acceptable internal consistency and good test-retest reliability (Tonigan, Connors, & Miller, 1996). Two items on the AAI specifically ask about the completion of steps in 12-step programs, the difference being that one item asks about the completion of specific steps only within the context of formal treatment and the second item inquires about completion of specific steps *regardless* of whether it was encouraged in formal treatment or within the context of attending community-based 12-step programs. We used the latter measure of steps completed as the direct measure of step work, and it was completed at each assessment point. Data from the 3-month assessment was used for the comparison with the GAATOR. In a previous study the test-retest correlation for the summed number of steps completed using this question was .98 (Tonigan, Miller, & Brown, 1997).

Form 90—The Form 90 (Miller, 1996) is an interviewer-administered questionnaire that includes information on substance use, help seeking, and health care utilization in the previous 90 days. The Form 90 has good test-retest reliability (Rice, 2007; Tonigan, Miller, & Brown, 1997). In a multi-site test-retest exercise, for example, intra-class correlation coefficients for proportion abstinent days, drinks per drinking day, and total alcohol consumption in standard drinks were .85, .71, and .82, respectively (Tonigan, Miller, & Brown, 1997).

Four measures of substance use were calculated using the Form 90: (1) *Percent days abstinent (PDA)* was defined as the number of days alcohol was not used in a period divided by the total number of days in that assessment period, (2) *Alcohol abstinence* was defined as complete abstinence from alcohol consumption during a period, (3) *Drinks per drinking day* was defined as the total number of standard drinks consumed in a period divided by the number of days on which alcohol was consumed (abstinent days excluded from the denominator), and (4) *Proportion days illicit drug use* was defined as the number of days an illicit drug was reported to have been used divided by the total number of days in an assessment period. *Proportion of days 12-step programs were attended* was deduced from the Form 90 by dividing the reported number of days 12-step meetings had been attended by the total number of days in an assessment period.

Results

Participant Characteristics

One individual did not complete the baseline GAATOR and, as a result, all sample descriptions and analyses are based upon $N = 129$. Participants (46.5% female) ranged in age from 19 to 64 years, with a mean age of 38.76 ($SD = 9.56$). Thirty-six percent of the sample was non-Hispanic White ($n = 47$), 33.3% was New Mexican/Spanish-American ($n = 43$), 15.5% was American Indian/Alaska Native ($n = 20$), and 10.9% was Mexican-

American ($n = 14$). At study intake, 57 were single (44.2%), 27 were divorced (20.9%), and 26 were married or cohabiting with a partner (20.2%). Participants ($N = 126$) reported a median yearly household income of \$10,000, range of \$0 to \$120,000. About 28% were employed. Half had a H.S. Diploma or GED (48.8%; $n = 63$), 18.6% ($n = 24$) had not completed high school, 22.5% had a trade school certificate or associate's degree ($n = 29$), and 10.1% had a bachelor's ($n = 8$) or master's ($n = 5$) degree.

At baseline participants consumed alcohol an average of 43.9% of the 90 days before recruitment ($SD = .31$) and consumed an average of 14.02 drinks per drinking day ($SD = 9.13$). Compared to baseline, participants reported 9.67 drinks per drinking day at 3 months ($SD = 9.13$; $d = .44$), 8.84 drinks per drinking day at 6 months ($SD = 10.78$; $d = .57$), and 8.60 drinks per drinking day at 9 months ($SD = 8.80$; $d = .60$). Abstinence rates (from both alcohol and illicit drugs) since the previous interview were respectively 24.0% ($n = 31$), 29.5% ($n = 38$), and 31.8% ($n = 41$) at the 3, 6, and 9-month follow-ups. On about 27% of the days before recruitment participants reported using illicit drugs ($SD = .37$).

The mean proportion of days of 12-Step meetings attended in the 90-day period before recruitment was .17 ($SD = .20$), which translates into about one meeting every five days. This figure was .25 ($SD = .31$) at 3 months, .23 ($SD = .29$) at 6 months, and .18 ($SD = .24$) at 9 months. At intake, 65.9% of the sample considered themselves to be members of AA. At the 3-, 6-, and 9-month follow-ups, 80%, 79%, and 76% of the participants reported attending a 12-Step meeting since the last interview (generally 90 days), respectively.

Correspondence between the AAI and GAATOR

At the three-month follow-up, the percentage of participants reporting step-work on the AAI ranged from a high of 32.2% for Step 1, to a low of 3.5% for Steps 10 and 11 (see Table 1). On the GAATOR, the percentage of participants reporting step-work ranged from 73.9% for Step 2, to 23.5% for Step 8. More participants tended to endorse working the earlier versus later steps, consistent with their limited exposure to 12-step programs. Cohen's kappa tests were conducted to test for agreement between percent of individuals endorsing 12-step work on the GAATOR versus the AAI. Correspondence between these two measures of step-work was for the most part poor. The highest kappa values were observed for Step 3 (kappa = .20) and 4 (kappa = .34), but using standard benchmarks these values are still considered only fair (Landis & Koch, 1977).

Factor Analysis of the GAATOR

Exploratory factor analysis (EFA) was used to investigate the underlying structure of GAATOR item responses separately at three and nine months. Prior to conducting the EFA, items within each step were summed such that variables entered into the EFA represented endorsement of steps 1 through 12 (see Table 1 for GAATOR items included in each step). Varimax rotation with alpha extraction yielded two interpretable factors with eigenvalues greater than 1.0 (Table 2) that replicated at intake, 3, 6, and 9 months.

At three months ($N = 113$), seven steps loaded significantly on the first factor (eigenvalue = 6.61; 31.3% of the item variance), while seven steps loaded significantly on the second factor (eigenvalue = 1.40; 29.2% of the item variance). The first factor included Steps 4, 5, 8, 9, 10, 11 and 12, steps that comprise concrete behavioral activities, such as taking inventory and making amends. This factor was labeled "Behavioral Step-Work." The second factor included steps 1, 2, 3, 6, 7, 11, and 12 – steps related to introspective AA practices, such as surrender and contact with a higher power. We labeled this factor, "Spiritual Step-Work." While steps 11 and 12 loaded significantly on both factors, step 11 loaded more

strongly on the Spiritual Step-Work factor, and step 12 loaded more strongly on the Behavioral Step-Work factor.

A similar factor structure was obtained using the GAATOR data collected at nine months ($N = 113$). Seven steps loaded significantly on the Behavioral Step-Work factor (eigenvalue = 6.76; 32.3% of the item variance), and seven steps loaded significantly on the Spiritual Step-Work factor (eigenvalue = 1.21; 27.4% of the item variance). The Behavioral Step-Work factor included Steps 4, 5, 8, 9, 10, 11 and 12, and the Spiritual Step-Work factor included Steps 1, 2, 3, 6, 7, 10 and 11. Unlike the three-month factor structure, at nine months step 12 only loaded significantly on the Behavioral Step-Work factor. Similar to three months, step 11 loaded significantly on both factors, but slightly more on the Spiritual Step-Work factor.

Combining the results of the three and nine-month EFA, subsequent analyses consider two GAATOR factors: Behavioral Step-Work (steps 4, 5, 8, 9, 10, and 12) and Spiritual Step-Work (steps 1, 2, 3, 6, 7, and 11). Across assessment points, alpha coefficients ranged from .87 to .92 for the Behavioral Step-Work scale and from .85 to .87 for the Spiritual Step-Work scale.

Trajectory of GAATOR Step-Work

Analyses were conducted using HLM 6.0 (Raudenbush, Bryk, Cheong, & Congdon, 2004). Time-varying covariates were added to the model based on their theoretical relationship to step-work, and included: proportion days 12-step attendance and formal treatment, having or not having a sponsor at a given follow-up (yes/no), and being encouraged by a treatment program to attend AA (yes/no). Time was centered at intake, and the four assessment points in the study were coded as 0, 3, 6, and 9 to provide information about net changes in step-work per month. Table 3 provides mean GAATOR subscale scores at each assessment. Cross-sectional bivariate correlations between the two GAATOR step scales collected at baseline, 3, 6, and 9- month follow-up were $r = .57$, $r = .67$, $r = .71$, and $r = .71$, respectively. Each of the two GAATOR scales had a possible range of 0-24.

Behavioral step-work—Changes in GAATOR Behavioral Step-Work scores over time were first examined with only the linear fixed effect of time and random intercept in the model. At baseline, the mean Behavioral Step-Work score was 6.67 ($SE = 0.43$), which was significantly different from zero ($p < .001$) indicating that participants had already begun working the 12 steps prior to study recruitment. The parameter estimate for the linear effect of time was non-significant, $b = 0.07$, $SE = 0.06$, $p = .22$, indicating that, on average, the endorsement and practice of Behavioral Step-Work did not change over time. When the time-varying covariates were added to the model (Table 4), only the main effect of having a sponsor was significant, $b = 1.72$, $SE = 0.50$, $p = .001$, suggesting that those who reported higher scores on Behavioral Step-Work also tended to report having a sponsor during that period. The main effect of treatment program attendance ($b = 3.21$), being encouraged by a treatment program to attend AA ($b = 1.09$), and 12-step attendance ($b = 2.20$) were non-significant. Adding this set of predictors to the model did not alter the non-significant effect of linear time on Behavioral Step-Work.

Spiritual step-work—Using the same two-step process, changes in Spiritual Step-Work over time were examined. Mean Spiritual Step-Work score at intake was 12.23 ($SE = 0.45$), which was significantly different from zero, $p < .001$, and which represented the practice of spiritual-focused steps prior to study recruitment. The linear effect of time was non-significant, $b = -0.10$, $SE = 0.06$, $p = .08$, meaning that, on average, self-reported Spiritual Step-Work did not change over time. Entering the time-varying covariates led to a significant linear effect of time, $b = -0.14$, $SE = 0.05$, $p = .007$ (Table 4). Within the context

of other 12-step related influences, then, Spiritual Step-Work was observed to decrease over time. However, the main effect of attending AA was significant, $b = 2.62$, $SE = 0.91$, $p = .005$, as was the main effect of attending treatment, $b = 5.13$, $SE = 1.64$, $p = .002$. Independent of time, therefore, higher rates of 12-step attendance and formal treatment were associated with more Spiritual Step-Work. This was not the case for the other time-varying covariates, encouragement to attend AA from a treatment program ($b = 0.85$), and having a sponsor ($b = 0.57$), which were non-significant covariates.

GAATOR Step-Work and Substance Use

Four lagged HLM analyses were conducted to assess the extent, if any, that proportion days 12-step attendance, endorsement of Behavioral and Spiritual Step-Work (collected at months 0, 3, and 6) and time (linear) uniquely predicted four measures of substance use (collected at months 3, 6, and 9). Results of these four conservative tests of the importance of step-work in predicting later substance use reductions can be seen in Table 5. First, substance use reductions were made between baseline and the 3-month interview, and these reductions were sustained (but did not continue to improve) throughout the remaining six months of the study (non-significant time effect). Second, proportion days 12-step attendance was a significant predictor of later abstinence regardless of whether abstinence was defined as a categorical or continuous outcome measure. Third, Spiritual Step-Work significantly predicted later abstinence when abstinence was defined as a continuous measure, but only at a trend level did this scale predict our categorical measure of abstinence ($p = .09$). And, fourth, Behavioral Step-Work failed to predict later reductions in any of the four measures of substance use.

Discussion

The prescribed beliefs and practices of AA are most succinctly stated in the 12 steps, and working the 12-steps is widely endorsed and considered by many to be the underlying reason accounting for AA-related benefit (Booth, 1987). Our findings indicate that directly asking early 12-step affiliates about the steps they have completed is a poor indicator of their endorsement and practice of the beliefs and practices embedded in the steps. Specifically, we found that participants' endorsement of steps completed was consistently lower than the endorsement of the beliefs and practice abstracted from each of the steps, significantly so for a majority of the 12 steps. It appears, then, that the endorsement of "completing" a step had a higher threshold among participants than the endorsement of practicing specific elements within a step. In this regard, efforts to understand the adoption of prescribed beliefs and practices of early AA affiliates on the basis of reported completion of steps is likely to systematically produce false negatives. At the same time, indirect measures such as the GAATOR may overestimate step-work completion to some degree by casting a wider net that includes general beliefs and practices that can occur both within AA and through other outlets (e.g., #19: "I have prayed and meditated").

A number of deductive strategies have been used to describe the function of the 12 steps, with one of the most frequently cited being the characterization of steps 1-3 as surrender steps, steps 4-9 as action steps, and steps 10-12 as maintenance steps (Tonigan & Rice, 2010). Exploratory factor analysis in this study indicated an alternative conceptualization of the 12 steps as practiced and endorsed by early AA affiliates; namely, behaviorally-oriented and spiritually-focused steps. Common to the behaviorally-oriented steps (steps 4, 5, 8, 9, 10, and 12) were prescriptions that called for overt behavior. Such behaviors included preparing a written self-inventory (step 4), sharing of an inventory with another (step 5), making a list of harmed persons (step 8), making amends to harmed persons (step 9), taking a daily inventory and promptly acknowledging wrong doing (step 10), and working with other alcoholics (step 12). In contrast, the spiritually-focused steps prescribed self-reflection

and practices intended to enhance spiritual growth: admitting powerlessness over alcohol (step 1), belief that a Higher Power could restore sanity (step 2), turning one's will over to the care of a Higher Power (step 3), humbly identifying shortcomings (step 6), asking God to remove these shortcomings (step 7), and practicing prayer and meditation (step 11).

The findings of our factor analysis are not qualitatively different than those reported by Tonigan et al. (2000). Specifically, they reported that the GAATOR consisted of three components, which they labeled *Higher Power*, *Self-Inventory*, and *Other Inventory*. The nine items within the Higher Power component identified by Tonigan et al. overlap considerably with those items in our spiritually-based scale. The key difference in findings resides in the manner in which prescribed behavioral practices in the steps were classified: Tonigan et al. reported that sample responses were partitioned between *self* and *other* behavioral practices while we found no such separation. We suspect that this nuance is explained by sample differences in the extent that participants attended and were engaged in AA. Well documented, for example, AA prescribed behaviors and beliefs are positively correlated during early AA affiliation, especially so when members are actively engaged in AA (e.g., Connors et al. 2001). Higher rates of AA engagement would logically result in more uniform rates of endorsement of both self and other behavioral practices whereas lower rates of AA engagement may result in differential endorsement of self-other prescribed behaviors. Described earlier, the conditions in which Tonigan et al. sampled AA exposed adults were not ideal and probably did not capture the full range of AA practices that is typically observed among adults attending AA.

This study offered a rare view of how the endorsement and practice of the 12 steps changed over a 9-month period among early AA affiliates. Aggregating endorsement of 12-step prescribed beliefs and practices into the behavioral and spiritually-based categories, for example, we found that the extent of endorsement of these categories did not change over time among participants if we ignored factors outside of the individual (e.g., treatment attendance; 12-step practices). In contrast, having a 12-step sponsor was associated with increased endorsement of the behaviorally-anchored practices and beliefs (but not spiritually-focused steps). This is probably the case because of steps 4 and 5. Important to note, our finding is at odds with Tonigan & Rice's (2010) report that having a sponsor at 3 months was unrelated to steps completed at 6 months. We suspect that the difference in findings may be the result of their combining spiritual and behavioral steps into a single scale as opposed to our distinguishing these types of steps.

Two factors were identified that influenced changes in early 12-step affiliates' endorsement of spiritually-focused step practices and beliefs. First, frequency of 12-step attendance was positively associated with increased endorsement of spiritually focused beliefs. Second, many participants reported concurrent formal treatment during study participation and such experiences also predicted increases in spiritually directed practices that are prescribed in the 12 steps. Treatment providers encourage the completion of Step 1, and this may account for some of the increase in spiritually directed practices.

Integrating the preceding discussion, our findings suggest that behavioral prescriptions in the 12 steps are realized through sponsorship while spiritual beliefs and practices in the prescribed 12 steps are acquired through 12-step meeting attendance and, for some, with concurrent formal treatment. This interpretation is wholly consistent with recent 12-step research. Specifically, four prospective studies have now reported that 12-step attendance is predictive of later increases in spirituality, with spirituality defined in a variety of ways (Kaskutas et al., 2003; Kelly et al., 2007; Zemore, 2007). Our study identifies a potential means by which such spiritual growth may occur in relation to 12-step meeting attendance,

namely the discussion in meetings of spiritual principles and practices outlined in steps 1, 2, 3, 6, 7, and 11.

While several studies have predicted substance use outcomes using composite measures of AA-related practices that include step-work (e.g., Johnson, Finney & Moos, 2006; Owen et al., 2003), and completion of individual steps has been associated with later sobriety (Carroll, 1993; Gilbert, 1991), this study offered a rare view of how aggregate step-work, controlling for frequency of 12-step attendance, relates to later substance use. Spiritual and Behavioral Step-Work did not predict drinks per drinking day or proportion days illicit drug use. However, Spiritual Step-Work did predict percent days abstinent, while Behavioral Step-Work did not predict either outcome. This is consistent with findings from other studies on the importance of spirituality for later abstinence (e.g., Kaskutas et al., 2003; Robinson, Krentzman, Webb, & Brower, 2011), and the current study advances this research by identifying what may be at the heart of mobilizing spiritual growth in 12-step programs: working the spiritually-related steps. Theoretically, benefits accumulated from behavioral step-work (e.g., taking inventory) may be less critical to sustained abstinence when individuals are faced with a high-risk relapse situation than benefits accrued from spiritual step-work (e.g., powerlessness, prayer), but this model requires further exploration.

Limitations

Several study limitations should be noted. First, by focusing exclusively on individuals with limited 12-step experience, we gained internal validity at the cost of generalizability. The effect of step-work may be cumulative, and the strength of the relationship between adoption of 12-step practices and abstinence may differ for individuals with extended 12-step experience. In this vein, behavioral step-work may have effects that were not identified in this study. Additionally, study participants were of primarily low SES and findings may be most applicable to individuals with similar backgrounds. On the other hand, this ethnically diverse, low SES sample can be seen as a strength because it extends research on 12-step programs beyond more frequently studied predominantly White middle class individuals. Second, the spiritual step-work precepts outlined in the GAATOR are grounded in Judeo-Christian beliefs. As a consequence, endorsement of GAATOR Spiritual Step-Work may also include spiritual beliefs acquired outside of the 12-step programs. Third, while our statistical analyses investigated the unique contributions of behavioral and spiritually-based practices expressed in the 12-steps, it is important to recognize that these two classes of practices do not occur in isolation from one another. The extent that one class of practices operates independently of the other is unknown but certainly requires attention. Fourth, our exploratory factor analyses (EFA) examined the underlying structure of 12 items corresponding to the 12 steps using a sample of 113 participants. Because this exceeded the frequently cited minimum ratio of 10 subjects per item in EFA, a larger sample would add further confidence to study findings. And, finally, recruitment efforts focused specifically on AA meetings, but our follow-up measure of 12-step attendance was broad and extended beyond AA to include attendance at any 12-step program. Our decision to capture the net influence of participation in 12-step programs could be improved upon in future research by recording attendance at specific 12-step programs. This would allow for both an assessment of the unique influences of 12-step sister programs, e.g., NA and CA, as well as deriving a cumulative effect by summing 12-step attendance across programs.

Summary and Conclusions

This study investigated the benefits and nuances of step-work, a key component of 12-step programs that has been widely espoused without clear empirical evidence to inform why step-work might be beneficial for substance use outcomes or how it can be mobilized. Our findings suggest that indirect measures of step-work such as the GAATOR produce fewer

false negatives than direct measures. They provide a new conceptual framework in which to consider step-work as behaviorally versus spiritually-oriented, highlight factors that contribute to increases in step-work, and demonstrate the importance of spiritually-focused step-work for later abstinence. Future studies should replicate our distinction between Behavioral and Spiritual Step-Work and utilize longer follow-ups to better understand the changing nature of step-work over time and its importance for substance use outcomes.

Acknowledgments

This research was supported by Grants R21AA016974 and T32AA018108-01A1 from the National Institute on Alcohol Abuse and Alcoholism. The views expressed are those of the authors and do not necessarily represent the views of NIAAA.

Appendix: The 12 Steps and corresponding GAATOR items

Appendix

The 12 Steps and corresponding GAATOR items

Step 1: We admitted we were powerless over alcohol – that our lives had become unmanageable

8. I have accepted the fact that I can never drink or use again.

13. I have realized that no matter what I do, things get worse when I drink or use.

Step 2: Came to believe that a Power greater than ourselves could restore us to sanity.

4. I have believed that my recovery could only come from a power greater than myself.

20. I have believed that awareness of my Higher Power is essential to my sobriety.

23. I knew I must have faith in order to keep my sobriety.

Step 3: Made a decision to turn our will and our lives over to the care of God *as we understood Him*.

1. I have turned my will and my life over to my Higher Power.

Step 4: Made a searching and fearless moral inventory of ourselves.

5. I have made a list of my resentments.

18. Writing down all my shortcomings has helped my sobriety.

Step 5: Admitted to God, to ourselves and to another human being the exact nature of our wrongdoings.

3. I have shared my personal inventory with another person.

14. I have shared my personal inventory with someone I trust.

Step 6: Were entirely ready to have God remove all these defects of character.

10. I have been ready to let my Higher Power remove my shortcomings.

12. I have found character defects which I am willing to give up.

Step 7: Humbly asked Him to remove our shortcomings.

6. I have recognized that the amount of serenity I have is a direct result of the amount of humility which I have.

15. I have asked my Higher Power to remove my defects of character.

Step 8: Made a list of all persons we had harmed, and became willing to make amends to them all.

16. I have made a list of people whom I had harmed.

21. I have discussed with another person how best to make amends.

Step 9: Made direct amends to such people whenever possible, except when to do so would injure them or others.

2. I have made direct amends to those whom I had harmed.

11. I have made indirect amends to those whom I had harmed, when direct amends were not possible.

Step 10: Continued to take personal inventory and when we wrong promptly admitted it.

7. I have taken a daily inventory of my behavior.

22. *I have watched for selfishness, dishonesty, resentment and fear, and when these appeared, I have asked my Higher Power to remove them.*

Step 11: Sought through prayer and meditation to improve our conscious contact with God, *as we understood Him*, praying only for knowledge of His will for us and the power to carry that out.

9. *I have told others about my spiritual experience.*

19. *I have prayed and meditated.*

Step 12: Having had a spiritual awakening as a result of these steps, we tried to carry this message to alcoholics, and to practice these principles in all our affairs.

17. *I have tried to practice the Twelve Steps in all my affairs.*

24. *I have shared my experiences, strength, and hope with others.*

Note. Response choices include Definitely False (0), False (0), True (1), and Definitely True (2). Responses for items listed as corresponding to the same step are summed to provide a total score for a particular step. Total scores for steps 4, 5, 8, 9, 10, and 12 are summed to form the Behavioral Step-Work scale (range = 0-24); total scores for steps 1, 2, 3, 6, 7, and 11 are summed to form the Spiritual Step-Work scale (range = 0-24). The GAATOR is available through the <http://casaa.unm.edu> website (select downloads; then assessment instruments).

References

- Bond J, Kaskutas LA, Weisner C. The persistent influence of social networks and Alcoholics Anonymous on abstinence. *Journal of Studies on Alcohol*. 2003; 64(4):579–588. [PubMed: 12921201]
- Booth L. Alcoholism and the fourth and fifth steps of Alcoholics Anonymous. *Journal of Psychoactive Drugs*. 1987; 19(3):269–274. [PubMed: 3694330]
- Brown HP, Peterson JH. Assessing spirituality in addiction treatment and follow-up: Development of the Brown-Peterson Recovery Progress Inventory (B-PRPI). *Alcoholism Treatment Quarterly*. 1991; 8(2):21–50. doi:10.1300/J020V08N02_03.
- Carroll S. Spirituality and purpose in life in alcoholism recovery. *Journal of Studies on Alcohol*. 1993; 54(3):297–301. [PubMed: 8487537]
- Connors GJ, Tonigan J, Miller WR. A longitudinal model of intake symptomatology, AA participation and outcome: Retrospective study of the Project MATCH outpatient and aftercare samples. *Journal of Studies on Alcohol*. 2001; 62(6):817–825. Retrieved from EBSCOhost. [PubMed: 11838919]
- Forcehimes AA, Tonigan JS. Self-efficacy as a factor in abstinence from alcohol/other drug abuse: A meta-analysis. *Alcoholism Treatment Quarterly*. 2008; 26(4):480–489. doi:10.1080/07347320802347145.
- Gilbert FS. Development of a “Steps Questionnaire.”. *Journal of Studies on Alcohol*. 1991; 52(4):353–360. [PubMed: 1875709]
- Groh DR, Jason LA, Keys CB. Social network variables in alcoholics anonymous: A literature review. *Clinical Psychology Review*. 2008; 28(3):430–450. doi:10.1016/j.cpr.2007.07.014. [PubMed: 17719158]
- Humphreys K, Kaskutas LA, Weisner C. The Alcoholics Anonymous Affiliation scale: Development, reliability, and norms for diverse treated and untreated populations. *Alcoholism: Clinical and Experimental Research*. 1998; 22(5):974–978.
- Johnson JE, Finney JW, Moos RH. End-of-treatment outcomes in cognitive-behavioral treatment and 12-step substance use treatment programs: Do they differ and do they predict 1-year outcomes? *Journal of Substance Abuse Treatment*. 2006; 31(1):41–50. doi:10.1016/j.jsat.2006.03.008. [PubMed: 16814009]
- Kaskutas L, 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. doi:10.1300/J020v21n01_01.
- Kelly JF, Dow SJ, Yeterian JD, Kahler CW. Can 12-step group participation strengthen and extend the benefits of adolescent addiction treatment? A prospective analysis. *Drug And Alcohol*

- Dependence. 2010; 110(1-2):117–125. doi:10.1016/j.drugalcdep.2010.02.019. [PubMed: 20338698]
- Kelly JF, Urbanowski KA, Hoepfner BB, Slaymaker VJ. Facilitating comprehensive assessment of 12-step experiences: A multidimensional measure of mutual-help activity. *Alcoholism Treatment Quarterly*. 2011; 29(3):181–203. [PubMed: 22081741]
- Kelly JF, Stout RL, Magill M, Tonigan J, 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. doi:10.1111/j.1530-0277.2010.01362.x.
- Krentzman AR, Robinson ER, Moore BC, Kelly JF, Laudet AB, White WL, Strobbe S. How Alcoholics Anonymous (AA) and Narcotics Anonymous (NA) work: Cross-disciplinary perspectives. *Alcoholism Treatment Quarterly*. 2011; 29(1):75–84. doi: 10.1080/07347324.2011.538318. [PubMed: 21785524]
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977; 33(1):159–174. [PubMed: 843571]
- Miller, WR. Form 90: A structured assessment interview for drinking and related behaviors (NIAAA Project MATCH Monograph Series, Vol. 5, NIH Publication No. 96-4004). National Institute on Alcohol Abuse and Alcoholism; Rockville, MD: 1996.
- Montgomery HA, Miller WR, Tonigan JS. Does Alcoholics Anonymous involvement predict treatment outcome? *Journal of Substance Abuse Treatment*. 1995; 12(4):241–246. doi: 10.1016/0740-5472(95)00018-Z. [PubMed: 8830150]
- Moos RH, Moos BS. Participation in Treatment and Alcoholics Anonymous: A 16-Year Follow-Up of Initially Untreated Individuals. *Journal of Clinical Psychology*. 2006; 62(6):735–750. doi:10.1002/jclp.20259. [PubMed: 16538654]
- Ouimette P, Finney JW, Gima K, Moos RH. A comparative evaluation of substance abuse treatment III. Examining mechanisms underlying patient-treatment matching hypotheses for 12-step and cognitive-behavioral treatments for substance abuse. *Alcoholism: Clinical and Experimental Research*. 1999; 23(3):545–551. doi:10.1097/00000374-199903000-00024.
- Owen PL, Slaymaker V, Torigan JS, McCrady BS, Epstein EE, Kaskutas LA, Miller WR. Participation in Alcoholics Anonymous: Intended and unintended change mechanisms. *Alcoholism: Clinical and Experimental Research*. 2003; 27(3):524–532. doi:10.1097/01.ALC.0000057941.57330.39.
- Pagano ME, Krentzman AR, Onder CC, Baryak JL, Murphy JL, Zywiak WH, Stout RL. Service to others in sobriety (SOS). *Alcoholism Treatment Quarterly*. 2010; 28(2):111–127. doi: 10.1080/07347321003656425. [PubMed: 21399711]
- Raudenbush, SW.; Bryk, AS.; Cheong, YF.; Congdon, RT, Jr.. HLM 6: Hierarchical Linear Modeling and Nonlinear Modeling. Scientific Software International; Lincolnwood, IL: 2004.
- Rice C. Retest reliability of self-reported daily drinking: Form 90. *Journal of Studies on Alcohol and Drugs*. 2007; 68(4):615–618. [PubMed: 17568968]
- Robinson ER, Krentzman AR, Webb JR, Brower KJ. Six-month changes in spirituality and religiousness predict drinking outcomes at nine months. *Journal of Studies on Alcohol and Drugs*. 2011; 72:660–668. [PubMed: 21683048]
- Roman, PM.; Johnson, JA. National Treatment Center Study Summary Report: Private Treatment Centers. Institute for Behavioral Research, University of Georgia; Athens, GA: 2004.
- Substance Abuse and Mental Health Services Administration. Office of Applied Studies. National Survey of Substance Abuse Treatment Services (N-SSATS): 2007. Data on Substance Abuse Treatment Facilities. Rockville, MD: 2008. OAS Series #S-44, DHHS Publication No. (SMA) 08-4343
- Tonigan JS, Beatty GK. 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, Rice SL. Is it beneficial to have an alcoholics anonymous sponsor? *Psychology of Addictive Behaviors*. 2010; 24(3):397–403. doi:10.1037/a0019013. [PubMed: 20853924]

- Tonigan J, Connors GJ, Miller WR. Alcoholics Anonymous Involvement (AAI) scale: Reliability and norms. *Psychology of Addictive Behaviors*. 1996; 10(2):75–80. doi:10.1037/0893-164X.10.2.75.
- 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.; Miller, WR.; Vick, D. Psychometric properties and stability of the General Alcoholics Anonymous Tools of Recovery (GAATOR, 2.1); Poster presented at the annual Research Society on Alcoholism conference; Denver, CO. Jun. 2000
- Withbrodt J, Kaskutas LA. Does diagnosis matter? Differential effects of 12-step participation and social networks on abstinence. *The American Journal of Drug and Alcohol Abuse*. 2005; 31(4): 685–707. doi:10.1081/ADA-68486. [PubMed: 16320441]
- Zemore SE. A role for spiritual change in the benefits of 12-step involvement. *Alcoholism: Clinical and Experimental Research*. 2007; 31(Suppl10):76S–79S. doi:10.1111/j.1530-0277.2007.00499.x.
- Zemore SE, Kaskutas L, Ammon LN. In 12-step groups, helping helps the helper. *Addiction*. 2004; 99(8):1015–1023. doi:10.1111/j.1360-0443.2004.00782.x. [PubMed: 15265098]

Table 1

Endorsement of Step-Work at 3 Months on the GAATOR and AAI

Step	GAATOR items ()	<u>Percent Endorsing Step-Work</u>	
		AAI (N = 115)	GAATOR (N = 115)
1	8, 13 (0.61)	32.2	40.9
2	4, 20, 23 (0.80)	24.3	73.9
3	1	31.3	65.2
4	5, 18 (0.84)	19.1	26.1
5	3, 14 (0.85)	8.7	36.8
6	10, 12 (0.63)	8.7	53.0
7	6, 15 (0.42)	7.8	48.7
8	16, 21 (0.70)	6.1	23.5
9	2, 11 (0.73)	4.3	27.0
10	7, 22 (0.48)	3.5	39.5
11	9, 19 (0.63)	3.5	38.3
12	17, 24 (0.64)	4.3	39.1

Note. AAI = Alcoholics Anonymous Inventory; GAATOR = General AA Tools of Recovery.

Table 2

Factor Loadings for GAATOR Exploratory Factor Analysis with Varimax Rotation (N=113)

Step	Three Months		Nine Months	
	Behavioral	Spiritual	Behavioral	Spiritual
8	.85	.16	.81	.26
4	.78	.32	.74	.33
5	.75	.24	.72	.22
9	.62	.27	.63	.28
10	.62	.42	.64	.50
12	.60	.52	.68	.41
2	.16	.77	.15	.80
6	.38	.77	.47	.72
3	.29	.75	.32	.61
11	.52	.63	.53	.61
1	.17	.59	.21	.54
7	.44	.56	.48	.72

Note. Factor loadings > .50 are in boldface. GAATOR = General Alcoholics Anonymous Tools of Recovery.

Table 3

Average GAATOR Scores Across Time

	Months			
	Intake	0 – 3	3 – 6	6 – 9
	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>
Behavioral Step-Work	6.60 (5.17)	6.69 (5.72)	7.77 (6.39)	6.96 (5.72)
Spiritual Step-Work	12.37 (5.50)	11.44 (5.74)	12.17 (5.41)	11.15 (5.74)

Note. At intake, 129 participants had scores for Behavioral Step-Work and 128 participants had scores for Spiritual Step-Work. At 3 months, 113 participants had scores for Behavioral Step-Work and 115 had scores for Spiritual Step-Work. At 6 and 9 months, 113 participants had scores for Behavioral and Spiritual Step-Work.

Table 4

Fixed Estimates for Final Models of Predictors of GAATOR Behavioral and Spiritual Step-Work

	Behavioral	Spiritual
Parameter	Estimate (SE)	Estimate (SE)
Intercept	5.02 (0.52) ***	10.84 (0.52) ***
Time	0.02 (0.05)	-0.14 (0.05) **
AA Attendance	2.20 (1.13)	2.62 (0.91) **
Sponsor	1.72 (0.50) **	0.57 (0.54)
Encouragement	1.09 (0.58)	0.85 (0.48)
Treatment Attendance	3.21 (2.18)	5.13 (1.64) **

Note. GAATOR = General Alcoholics Anonymous Tools of Recovery.

**
 $p < .01$.

 $p < .001$.

Table 5

Fixed Estimates (p values) for Lagged Models Predicting Four Substance Use Measures

<u>Predictors (Months 0, 3, 6)</u>	<u>Dependent Measures (Months 3, 6, and 9)</u>			
	DPDD	PctAbst	Illicit drug	Abst
Time	-.32 (.44)	.00 (.83)	-.02 (.19)	.09 (.57)
Proportion days 12-step Attended	-5.05 (.06)	.13 (.05)	-.04 (.30)	2.62 (.001)
Behavioral step-work	-.08 (.58)	-.01 (.17)	.001 (.76)	-.02 (.62)
Spiritual step-work	.10 (.56)	.01 (.05)	-.01 (.25)	.09 (.09)

Note. Boldface indicates $p < .05$. DPDD = drinks per drinking day; PctAbst = percent days abstinent from alcohol; Abst = alcohol abstinence.