

## A Randomized Trial Investigating Training in Motivational Interviewing for Behavioral Health Providers

Theresa B. Moyers and Jennifer K. Manuel

*University of New Mexico, Albuquerque, USA*

Paul G. Wilson

*Elmendorf Air Force Base, USA*

Stacey M. L. Hendrickson

*University of New Mexico, Albuquerque, USA*

Wayne Talcott

*Lackland Air Force Base, USA*

Peter Durand

*North Florida/South Georgia Veterans Health System*

**Abstract.** Evidence indicates that workshop training, personalized feedback, and individual consultation can increase competence in motivational interviewing (MI) among highly motivated and skilled substance abuse counselors. Little is known, however, about the translational value of these training strategies for counselors with fewer counseling skills and less stated motivation to learn MI. This study presents evidence from a randomized, controlled trial of 129 behavioral health providers assigned to receive workshop training and enrichments to learn MI. A diverse group of Air Force behavioral health providers working in substance abuse treatment programs were trained in MI and subsequently observed in clinical sessions at 4, 8 and 12 months after training. Results indicate that training was effective in increasing the skill level of these clinicians; however, these gains had decreased by the 4-month follow-up point. Training enrichments in the form of personalized feedback and consultation phone calls did not have an expected, additive effect on clinician skill level. The results of this

---

Reprint requests to Theresa B. Moyers, Department of Psychology, Center on Alcoholism, Substance Abuse and Addictions, University of New Mexico, Albuquerque, N.M., 87131-1161, USA. E-mail: tmoyers@unm.edu

© 2007 British Association for Behavioural and Cognitive Psychotherapies

study lend support to the hypothesis that a greater investment of resources and incentives may be necessary to achieve gains in MI skills for counselors with relatively lower baseline skills than those commonly participating in research studies.

*Keywords:* Motivational interviewing, dissemination, psychotherapy training, alcohol misuse.

## Introduction

Decades of research to illuminate strategies for intervening in substance abuse have yielded an array of promising treatments; however, a large gap exists between research and standard clinical practice (Miller, Sorenson, Selzer and Brigham, 2006). Current challenges in the substance abuse treatment field focus on translating these empirically-based treatments to front-line clinical settings. In particular, there is a need to identify which practitioners will be a good match for learning empirically-based treatments and whether research-based training methods will be feasible in real-world settings (Carroll and Rounsaville, 2003; Najavits, Crits-Christoph and Dierberger, 2000). Such translational research is sparse, despite the growing demand on providers and agencies to implement evidence-based treatments.

One empirically supported treatment now in widespread use is motivational interviewing (MI), a client-centered, directive treatment that focuses on helping clients resolve ambivalence about making adaptive lifestyle changes (Miller and Rollnick, 2002). Research supports the use of MI for addictive behaviors as well as promoting specific healthy behaviors, with effect sizes in outcome studies typically approaching 0.7, declining to 0.3 after one year (Hettema, Steele and Miller, 2005).

Dissemination studies with MI have so far focused on the conditions necessary for training providers to deliver this complex psychotherapy. Several studies have documented the ability of providers to gain knowledge *about* MI (Miller and Mount, 2001; Rubell, Sobell and Miller, 2000; Shafer, Rhode and Chong, 2004; Walters, Matson, Baer and Ziedonis, 2005). Such knowledge, though a necessary first step in learning a new therapeutic method, has minimal relationship to subsequent client outcomes (Carroll et al., 2006; Tappin et al., 2000).

Studies that more directly examine counselor skills after training in MI have typically used a format in which participants are assessed via behavioral observation both before and after workshop training. When the focus is on demonstrated skills rather than knowledge acquisition, results are mixed. Project EMMEE (Miller, Yahne, Moyers, Martinez and Pirritano, 2004) found large skill gains among a group of substance abuse treatment providers after training in MI, but other studies show much more modest improvements (Baer et al., 2004; Miller and Mount, 2001; Schoener, Madeja, Henderson, Ondersma and Janisse, 2006; Shafer et al., 2004). These results are confounded, however, by therapist motivation and skill level. In general, studies using highly motivated and skilled therapist samples (Baer et al., 2004; Miller et al., 2004) show larger training gains than those conducted in more realistic, front-line clinical settings where clinicians typically exhibit lower pre-training counseling skills and less expressed motivation for training (Schoener et al., 2006). Previous MI training studies have also been characterized by a lack of diversity within the ethnic background of the participants, limiting generalizability to common clinical settings. To increase the generalizability of MI training research, studies are needed that incorporate providers who are most likely to be found in front-line treatment clinics: ethnically diverse, with modest desire to learn an empirically-based treatment and highly variable general counseling skills.

The current study focuses on the dissemination of MI among behavioral health providers in the United States Air Force. Diverse in experience, professional affiliation and previous training in therapeutic interventions, this group represents a challenging target for dissemination of a complex clinical method. Within the Air Force, substance abuse treatment providers have not necessarily received formal training preparing them for counseling and may have been serving in an entirely different job category immediately prior to an assignment as a substance abuse counselor. As with their civilian counterparts, workshop and on-the-job training may comprise a vital part of their education (McLellan, Carise and Kleber, 2003; Walters et al., 2005). Air Force counselors, like nonmilitary substance abuse counselors, usually work with clients who have been referred by others for behavior that violates system norms.

This study used a randomized controlled design to assess the impact of MI workshop training in this group of behavioral health providers who were diverse in motivation for training, baseline skill and ethnicity.

## Methods

### *Participants*

Our intended participant pool was comprised of behavioral health providers working directly with substance abuse clients in the Air Force (AF). Individuals with a degree in psychology, medicine, social work, nursing, counseling or a certified alcohol and drug abuse counselor (CADAC) or other substance abuse treatment certification were eligible for enrollment.

We obtained the approval of three internal review boards at the University of New Mexico, the US Army Medical Research and Material Command at Fort Detrick, and Wilford Hall Medical Center. Once approval from all three oversight committees had been obtained, we advertised information about the study through the Air Force Alcohol and Drug Prevention Teams (ADAPT). We also advertised information about this study at the 2001 AF Worldwide Conference in San Diego. Potential participants contacted the PI directly, or obtained further information through the AF Surgeon General's Office.

### *Design overview and description of training groups*

Prior to randomization, participants signed an informed consent document and submitted a baseline audiotaped work sample. Participants were instructed to provide any tape of themselves with an actual client that demonstrated their typical clinical work. Once participants were enrolled in the study, they were randomly assigned to one of three conditions: Workshop Training (W), Workshop Training plus Training Enrichments (E) and Self-Directed Training (S). Workshop training was conducted by the first two authors in participant groups of 12–40 participants. All participants received a copy of *Motivational Interviewing* (Miller and Rollnick, 2002) and the Professional Learning Series videotapes (Miller, Rollnick and Moyers, 1998).

*Workshop Training (W).* Participants in the W group completed a 2-day motivational interviewing workshop held in Santa Fe, New Mexico or San Diego, California. The workshop followed the format used in a previous study (Miller et al., 2004), which focused on a learning-to-learn model for MI. That is, participants were introduced to MI concepts and taught methods for improving MI practice based on client responses. Immediately following the training,

participants completed a role-play with a standardized client, yielding the post-training work sample.

*Workshop Training plus Training Enrichments (E).* Participants in this group attended the same (conjoint) MI training workshops as the W group and submitted a post-training work sample. In addition, they received feedback on their post-training work sample, as well as up to six consultation phone calls with the goal of improving and extending the workshop training. Consultation calls followed exactly the format of the Miller et al. study (2004) and included role-plays and specific topics to enhance MI skills. Consult calls lasted 30 minutes and were intended to occur roughly 2 weeks apart.

*Self-Directed Training (S).* This group received the book and videotapes, but workshop training was delayed until the end of the study period in order to estimate the effect of self-directed learning. Within the study period, when the W and E groups were receiving workshop training and submitting follow-up work samples, the S group similarly submitted work samples, so that we could estimate the impact of self-directed training in the absence of any formal instruction. These work samples followed the same time-line of the W and E groups, but the workshops for the S participants had not yet occurred. Like the other conditions, this group provided a tape of themselves using MI with an actor at the conclusion of the workshop training (post-training tape). This was the last tape submitted by this group.

#### *Follow-up work samples*

For those in the W and E groups, participants were asked to submit audiotaped work samples of themselves using MI with an actual client in their work setting after completing workshop training and returning to their home work settings. These work samples were to be collected at 4, 8 and 12 months after workshop training. Tapes for the first follow-up period were submitted, on average, 5.4 ( $SD = 2.1$ ) months after training (this includes the Workshop and Enhanced groups only). Tapes for the second follow-up period were submitted, on average, 9.6 ( $SD = 2.3$ ) months after training, and tapes for the third follow-up period were submitted on average 11.5 ( $SD = 2.5$ ) months after training. Members of the S group also submitted tapes at these time periods despite not having yet completed workshop training.

#### *Description of the MITI*

All audiotaped work samples, including the baseline and post training samples, were evaluated using the Motivational Interviewing Treatment Integrity (MITI) scale (Moyers, Martin, Manuel and Miller, 2003), and this yielded the main outcome variables for the study. The MITI assesses clinician competence in MI using global evaluations of empathy and MI spirit as well as frequency counts of specific behaviors. MITI codes are mutually exclusive, such that each clinician utterance is only given one code. The MITI shows acceptable internal consistency and interrater reliability in evaluating specific clinician skills relevant to the use of MI (Moyers, Martin, Manuel, Hendrickson and Miller, 2005). Behavior counts from the MITI were aggregated to form summary measures (defined in Table 1) similar to those used in other studies assessing MI competence (Baer et al., 2004; Miller et al., 2004).

**Table 1.** Calculation of MITI summary measures

MITI summary measure	Formula
Ratio Reflections to Questions (R:Q)	$= \frac{\text{simple} + \text{complex reflections}}{\text{open} + \text{closed questions}}$
% Open Questions (%OQ)	$= \frac{\text{open questions}}{\text{open} + \text{closed questions}}$
% Complex Reflections (%CR)	$= \frac{\text{complex reflections}}{\text{simple} + \text{complex reflections}}$
% MI-Adherent (%MIA)	$= \frac{\text{MI Adherent behaviors}}{\text{MI Adherent} + \text{MI Non-adherent behaviors}}$

### *Process for review of tapes*

Over a 14-month period (January 2004 – March 2005) four coders reviewed all audiotaped work samples, including the baseline and post training samples, with the MITI. Coders were masked to the condition of the participant, including group assignment and assessment point. Tape coding did not begin until both baseline and post training tapes were available to preserve this blind.

### *Training and supervision of coders*

Four undergraduates from the University of New Mexico were initially trained to use the MITI with 50 hours of instruction and group practice. A series of graded learning tasks focusing on individual elements of the MITI were completed as coders progressed in learning this system. Competence at one level was achieved before progressing to more complex elements. For example, coders began by learning to differentiate open and closed questions before learning to differentiate simple from complex reflections. Proficiency was determined by comparing coder scores with gold standard examples. Only when coders had achieved acceptable reliability were they assigned tapes from the data pool. Throughout the tenure of the project, coders attended weekly meetings in which tapes were coded by the entire group to prevent drift from the coding manual and enhance interrater reliability. A random sample of 10% of tapes was double-coded as the initial coding of each 100 tapes was completed.

## **Results**

### *Description of participants*

In a 3-year period we randomized 129 substance abuse providers from 54 different Air Force bases. Demographic characteristics for this sample are listed in Table 2. Men (48.7%) and women were equally represented in this sample, with a mean age of 36.3 ( $SD = 8.9$ ). These providers had an average of 11.1 ( $SD = 7.0$ ) years experience in overall mental health counseling, with an average of 5.9 ( $SD = 6.2$ ) years of specific experience in the treatment of substance abuse. There were no significant differences between the treatment conditions on ethnicity, years of post-secondary education, or years in overall mental health counseling or treatment of substance abuse. However, there was a significant difference in the percentage

**Table 2.** Demographic characteristics of study participants

Characteristic	W	E	S
<i>N</i>	47	45	37
% female	49	35	77
% White/non-Hispanic	49	65	54
% master's or more	50	65	52
% doctorate	18	14	22
Postsecondary education			
<i>M</i> (years)	5.2	6.4	5.4
<i>SD</i>	3.4	3.7	4.1
Counseling experience			
<i>M</i> (years)	10.9	11.1	11.2
<i>SD</i>	7.6	5.1	8.4
Substance abuse counseling			
<i>M</i> (years)	6.4	4.9	6.5
<i>SD</i>	7.2	3.9	7.3

*Note:* W = workshop only; E = workshop plus training enrichments; S = self-directed.

of females within the Self-directed group compared to the other two treatment groups ( $\chi^2(2, N = 117) = 12.09, p = .002$ ).

Compared to the landmark Miller et al. (2004) training study, our participants were significantly less educated overall ( $t(246) = -3.60, p < .001$ ), had fewer years of experience in counseling ( $t(250) = -3.69, p < .001$ ), as well as less specific experience with substance abuse clients ( $t(248) = -5.45, p < .001$ ). The participants in our study were younger ( $t(158) = -8.70, p < .001$ ) and reported less interest in obtaining MI training ( $t(155) = -3.51, p = .001$ ) at the outset. Our participants were substantially more ethnically diverse, with 55.7% describing themselves as White, non-Hispanic compared to 88.6% in the Miller et al. (2004) study.

#### *Participant attrition and follow-up audiotaped work samples*

Of the 129 participants who were randomized and provided a baseline audiotaped work sample, 116 (90%) subsequently attended one of seven workshops in motivational interviewing and provided an audible post training tape. At the first follow-up point, 88 (68%) provided an audible work sample tape, with 75 (58%) and 49 (38%) providing audible work samples at the second and third follow-up periods. There were significant differences across training conditions in the likelihood of submitting audiotaped work samples ( $\chi^2(8) = 15.87, p < .05$ ), with the S participants more likely to submit follow-up tapes. The attrition rates for the trained participants (W and E) were 49% and 32% for the 8 and 12 month assessments respectively. We suspect that higher compliance rates in the S group may have been influenced by the fact that workshop training was only available to them after all audiotaped work samples had been submitted. Therefore, participants in the Self-directed group may have had a greater incentive to submit work samples. Because of the attrition rate among the trained providers at the 8 and 12 month follow-up points, we omitted 8 and 12 month data from any further analyses.

**Table 3.** Intra-class correlation coefficients at the 4 double-coding time points

Measure	Group 1	Group 2	Group 3	Group 4	All
Empathy	.419	.587	.435	.233	.431
MI Spirit	.517	.690	.513	.551	.608
% MI Adherent	.344	.806	.791	.921	.824
% Open Questions	.967	.935	.952	.915	.937
% Complex Reflections	.795	.301	.564	.784	.688
Ratio Reflections to Questions	.897	.942	.947	.938	.932

*Note:* Each group consisted of a set of 10 double coded audiotaped work samples randomly selected from the first, second, third, or fourth 100 coded work samples. Each group  $n = 10$ ; overall  $N = 40$ .

Participants in the Enhanced condition were eligible to receive up to six consult calls. Of the 45 participants in this condition, 16 (36%) did not complete any consult calls. Of the 29 (64%) that completed at least one consult call, three completed one call, three completed two calls, five completed three calls, three completed four calls, six completed five consult calls, and nine completed all six consult calls. Although all consult calls were to be completed prior to the 4-month assessment, 10 (34%) were submitted after this assessment. All participants in the E condition, even those not completing a single consult call, received personalized feedback regarding the standardized coding of their post training tape.

#### *Reliability of the dependent measures*

Intraclass correlations (ICCs) were calculated to assess reliability of the coding of the dependent measures. For each group of 100 tapes coded, we immediately double-coded a randomly selected sample of 10 tapes, resulting in four independent reliability estimates. This allowed us to calculate reliability estimates throughout the entire life of the project. In addition, we were able to identify and correct areas of poor coder reliability as the project progressed.

According to Cicchetti's (1994) system for evaluating intraclass correlations, those below 0.40 are considered to be poor, 0.40–0.59 are fair, 0.60–0.74 are good, and 0.75–1.00 are excellent. In general, the overall reliability estimates for our dependent measures were in the good to excellent range, although isolated examples of unacceptable reliability did occur as coding progressed. The ICC values for the dependent measures at all four double-coding time points are included in Table 3.

#### *Effectiveness of training*

Summary values for the primary outcome measures for each of the three randomized groups are displayed in Table 4. To assess whether workshop training produced improvements in MI skills, we conducted a repeated measures multivariate analysis of variance (MANOVA) on the difference between baseline and post training MITI scores, collapsing across training conditions. The Self-directed participants were included in this analysis. There was a significant difference between the two time points (Wilks' lambda,  $F(7, 106) = 82.39, p < .001, \eta_p^2 = .85$ ). Univariate tests reflected an improvement on all measures at the post training time period

**Table 4.** Means (*SD*) of primary outcome variables for each treatment group

Measure	W	E	S
Empathy			
Baseline	3.57 (0.93)	3.80 (1.22)	3.78 (1.16)
Post training	4.70 (0.84)	5.35 (0.58)	5.07 (1.16)
1 <sup>st</sup> follow-up	4.30 (1.56)	4.28 (1.33)	3.38 (1.18)
MI Spirit			
Baseline	2.60 (0.77)	3.09 (1.13)	3.19 (1.20)
Post training	4.46 (1.13)	4.83 (0.84)	5.31 (0.85)
1 <sup>st</sup> follow-up	3.93 (1.48)	4.00 (1.41)	2.97 (1.05)
MI Adherent Behaviors			
Baseline	1.79 (2.13)	2.02 (1.74)	2.00 (2.44)
Post training	4.63 (3.14)	4.93 (2.45)	3.14 (2.26)
1 <sup>st</sup> follow-up	1.43 (1.55)	2.03 (1.70)	1.90 (1.88)
MI Non-adherent Behaviors			
Baseline	2.34 (3.00)	1.73 (2.16)	0.81 (1.68)
Post-training	0.85 (1.99)	0.45 (0.88)	0.21 (0.41)
1 <sup>st</sup> follow-up	0.90 (1.42)	1.17 (1.73)	1.14 (1.53)
% Open Questions			
Baseline	21.13 (14.36)	18.97 (11.71)	18.79 (12.30)
Post training	44.33 (16.15)	42.45 (18.47)	51.63 (17.97)
1 <sup>st</sup> follow-up	31.94 (18.35)	28.03 (18.48)	21.29 (13.47)
% Complex Reflections			
Baseline	17.21 (17.01)	22.07 (21.14)	16.27 (16.32)
Post training	29.83 (13.92)	28.17 (13.03)	48.92 (14.14)
1 <sup>st</sup> follow-up	41.31 (26.25)	46.77 (19.49)	37.54 (23.31)
Reflection:Question ratio			
Baseline	0.35 (0.35)	0.39 (0.30)	0.35 (0.25)
Post training	1.13 (0.75)	1.32 (0.87)	1.50 (1.28)
1 <sup>st</sup> follow-up	0.77 (0.85)	1.05 (1.19)	0.49 (0.59)

*Note:* The post training audiotaped work sample was conducted with an actor/client and was collected immediately after the workshop training. The self-directed group (S) did not attend the workshop training until after the 3<sup>rd</sup> follow-up assessment. W = workshop only; E = workshop plus training enrichments; MI = motivational interviewing.

(all  $ps < .001$ ). Effect sizes for these observed gains in MI competence from baseline to post training were in the medium to large range (see Table 5).

To investigate the clinical, versus statistical, impact of improvements after workshop training we examined the average ratings for our clinicians and compared them to *a priori* recommended levels for MI practice (Moyers et al., 2003). Within the MITI coding system, clinical proficiency for the global measures of empathy and MI spirit is defined as a rating of five or above. For behavior counts, proficiency is demonstrated with values of percent open questions at least 50%, percent complex reflections at least 40%, the ratio of reflections to questions to be one or greater, and the percent MI adherent speech 90%. Mean post training ratings for participants in this study met or exceeded three of these threshold values (empathy, percent MI adherent and the ratio between reflections and questions). The percentage of clinicians meeting or exceeding



**Table 5.** Effect sizes (Cohen's *d*) for gains in MI competence from baseline to post training

Measure	All groups	W	E	S
Empathy	1.39	1.24	1.60	1.40
MI Spirit	1.90	1.93	1.68	2.47
% MI Adherent	1.14	1.76	1.09	0.53
% Open Questions	1.73	1.53	1.46	2.65
% Complex Reflections	0.91	0.77	0.34	2.48
Ratio Reflections to Questions	1.33	1.37	1.41	1.28

Note: W = workshop only; E = workshop plus training enrichments; S = self-directed.

**Table 6.** Tukey's HSD follow-up comparisons between treatment groups at post assessment

Measure	Group comparison	Mean difference	Significance ( <i>p</i> )
Empathy	S v. W	0.49	.037
	E v. W	0.61	.003
MI Spirit	S v. W	0.93	< .001
	S v. E	0.61	.032
% Open Questions	S v. E	0.11	.044
% Complex Reflections	S v. W	0.19	< .001
	S v. E	0.22	< .001

Note: S = self-directed; W = workshop only; E = workshop plus training enrichments.

all six criteria for MI proficiency at the post training point was 4.3% for the W group, 7.5% for the E group and 10.3% for the S group.

To explore differences in training gains among the groups, we conducted a MANOVA comparing the three training groups (W, E and S) with the six summary measures as dependent variables at the post training point. We found a significant overall difference (Wilks' lambda,  $F(12, 204) = 5.94, p < .001, \eta_p^2 = .26$ ). Follow-up contrasts demonstrating differences among the groups are found in Table 6.

To assess the possible influence of the gender imbalance in the SD group on study results, we compared men and women within each group on each outcome variable and found no significant gender differences.

We conducted a repeated measures MANOVA comparing the baseline scores to the scores at 4 months after the workshop was completed. We did not include the S group in this analysis as this group did not submit a true 4 month post training work sample. The comparison between the baseline and first follow-up assessment for the Workshop and Enhanced groups was significant ( $F(7, 52) = 11.21, p < .001, \eta_p^2 = .60$ ). Univariate comparisons showed improvement for all the follow-up variables of empathy ( $p = .009$ ), MI spirit ( $p < .001$ ), MI non-adherent speech ( $p = .012$ ), percent open questions ( $p < .001$ ), percent complex

reflections ( $p < .001$ ), and the ratio of reflections to questions ( $p = .001$ ). The only variable that did not exhibit sustained improvement was the frequency of MI adherent speech.

#### *Differences in skill across training conditions*

A multivariate analysis of covariance (MANCOVA) was performed to determine differences in MI skills between the Self-directed group and the groups that had received training (Workshop only and Enhanced) at the first follow-up point, with baseline MI skills as covariates. At the first follow-up point, the training groups (Workshop only and Enhanced) had received all the training components and the Self-directed group had only received the reading materials and videotapes. The overall F indicated a significant difference between the groups,  $F(7, 72) = 2.40, p = .029, \eta_p^2 = .189$ . Follow-up tests with univariate analyses of covariance (ANCOVAs) found significant differences between the Self-directed and trained groups in the dependent measures of empathy ( $p = .002, \eta_p^2 = .11$ ), and MI spirit ( $p = .001, \eta_p^2 = .13$ ) indicating an increase in MI skill for the trained groups in these areas. The measures of MI adherent speech, MI non-adherent speech, percent open questions, percent complex reflections, and the ratio of reflections to questions were not significant.

#### *Impact of training enrichments on workshop training in MI*

For this analysis we included only the Workshop and Enhanced conditions since the Self-directed group had not yet been trained. We conducted a MANCOVA comparing these two groups at the first follow-up with baseline values as covariates. Outcome measures were MI adherent behaviors, MI non-adherent behaviors, percent open questions, percent complex reflection, and ratio of reflections to questions, as well as global measures of empathy and MI spirit. This analysis failed to show a significant difference between the groups ( $F(7, 44) = 1.05, p = .41, \eta_p^2 = .14$ ).

#### *Differences in skill acquisition across military bases*

Some military bases had more than one behavioral health provider participating in the study. We conducted a MANCOVA to examine whether providers located on bases with other participating providers had better skill acquisition because of the extra peer support. Skill acquisition was measured at the first follow-up with the baseline MITI scores serving as the covariates. There was no significant difference in the skill levels at the first follow-up between sole providers and those located on a base with other participating therapists (Hotelling's Trace,  $F(6, 64) = 1.23, p = .30, \eta_p^2 = .10$ ).

## **Discussion**

The results of this randomized trial of training methods for MI lend support to the notion that this complex therapeutic treatment can be disseminated within diverse and demanding clinical settings. MI training in this study produced gains in skills among providers with highly variable professional and educational experience prior to instruction. Our participants showed gains in MI competence very similar to those of the substance abuse clinicians from National

Institute of Drug Abuse (NIDA) Clinical Trials Network (CTN) in the Smith et al. (2007) study.

Although therapeutic skills in this study did increase as a result of training, the gains were not as large as those observed in other clinical samples, and did not attain the level required for clinical trials (Miller et al., 2004). This sample of AF behavioral health providers gained substantially less skill than previous groups studied (Baer et al., 2004; Miller et al., 2004), despite nearly identical training and enrichment strategies. This finding is not surprising given the initial lower education, experience and baseline skill level of our participants compared to that observed in other MI training studies (Baer et al., 2004; Miller et al., 2004). These data lend support to the commonly expressed concern that a foundation of basic clinical skills may be needed to achieve proficiency in MI, especially given the relatively minor investment of time in these training paradigms (2-day workshop and six consult calls).

A decay of skills after initial training is common in clinical training studies (Walters et al., 2005) and our participants were no exception to this rule. Although clinicians in this sample demonstrated increased MI skills after workshop training, those gains showed a decline by the 4-month follow-up when compared to their performance immediately post training. This effect may have been exacerbated by using standardized patients to obtain work samples at post training, perhaps resulting in inflated scores and therefore larger differences at the first follow-up point. Though not unexpected, this decay of skills is troubling in the context of treatment systems attempting to implement complex empirically-based methods such as MI. It is likely that ongoing supervision and review of practice will be needed to maintain competence in newly trained practitioners, once workshop training has occurred. For example, Smith et al. (2007) utilized direct observation of therapy sessions as well as between session phone supervision resulting in 7.5 hours of post training enrichment and found less decay in obtained skills than clinicians in our study. Carroll et al. (2006) also have addressed this issue by following workshop training with intensive on-site supervision and review of session tapes to sustain competence for MI clinical trials within the NIDA CTN.

The impact of enrichments in our study was a surprise. We had expected to see improved performance in the group that received feedback and consult calls at the 4-month follow-up point, but we did not see it. Consult calls in this study did not improve skill level over a workshop alone, in contrast to the Miller et al. (2004) findings, despite the fact that every participant in the Enhanced condition received personalized coded feedback of their post training work sample within one month of the training workshop. This finding should be viewed with caution, however, since participants in the Enhanced condition completed only 44% of their available consult calls, and 34% of those occurred after the first follow-up. It is possible that enrichments would have shown a greater impact if participants had completed a larger percentage of the intended training enhancements. It is not clear why participants in this study were less willing to invest time in enriching their training through follow-up consult calls. Difficulty in accessing providers within military bases, differences in time zones between the consultant and the participants (many participants were located overseas), and the lack of direct incentives to increase their competence in their work settings may have contributed to this lack of engagement in consult calls. The most commonly cited reason for failing to complete consult calls was the lack of time or pressing clinical needs at the time the call was scheduled. Given the similarity between our population and those likely to be found in most treatment settings, it is possible that the level of compliance with training enrichments

found in this study offers a realistic estimate of what clinicians would be willing to do in other challenging clinical environments.

Another unexpected finding in our study was the superior performance of the Self-directed group (S) once they received workshop training. Assignment to this condition meant that participants had to wait the entire duration of the training phase of the study (up to 18 months) before receiving the MI workshop, while their counterparts in the other two conditions received it immediately. This meant that the S group had a longer period of time to use the self-study materials such as the book and training tapes and this may have accounted for their superior performance when compared to the Workshop (W) and Workshop-plus-Enrichment (E) groups at the post training assessment. Anecdotally, we received feedback from the S group about their disappointment at having to wait for the training workshop, which entailed a trip away from their duty stations and the aspect of a vacation for at least some participants. It is possible that having to wait for this “reward” caused the S group to be more diligent in their use of the study materials during that waiting period and more engaged in the training process during the workshop. Questioning of the participants to gather information about this hypothesis (and other questions about reasons for lack of compliance in returning tapes) was disallowed by the institutional review boards overseeing the study in favor of protecting the participants from coercion and potential repercussions from supervisors if their lack of compliance had become known, especially since some of the participants shared a work environment with others in the study.

Several limitations in this study are apparent. First, work samples in this study were self-selected since we asked participants to send us tapes showing us their “best effort” at using MI. This sampling method had the intentional advantage of allowing us to see whether participants *could* employ the skills we had hoped to teach, but does not provide any information about whether they are employed within the work setting more broadly. Second, similar to other substance abuse training studies, we suffered substantial attrition of work sample tapes after training. At month four, only 68% of our participants returned a follow-up work sample and by the one year follow-up we had only a 38% rate of return.

In this regard, our study is not unique. In general, poor follow-up rates are at least as high in studies for therapists receiving training as for clients receiving treatments. It is possible that incentives such as those that are used to increase client participation in treatment studies (compensation, lotteries, staggered payments for later samples) will be needed to ensure representative samples, and therefore increased internal validity, for dissemination studies with clinicians as the focus. It is worth noting that participants in this study returned tapes at a rate nearly identical to the Miller et al. (2004) sample, despite the fact that 21% of them were deployed to combat zones during the tenure of this project and that they were not permitted to receive incentive payments at any time. Further, it is clear that longer follow-ups to assess durability of clinical training are uncommon. Ours is only the second study that followed clinicians more than 4 months after initial training in motivational interviewing.

A potential confound in our findings is the prospect of cross-fertilization among therapists who were trained at the same site. It is possible that those clinicians trained at sites where peers were also enrolled in this study would have an advantage because they would be able to receive direction or input from each other. Our original design specified that bases would be randomized to each of the conditions, thereby preventing cross-fertilization across the groups (i.e. a participant from the Enhanced condition discussing these enrichments with a participant from the Workshop only condition). This plan was quickly abandoned in the face of high

mobility in our participants (meaning that we could not maintain the purity of base assignment once participants moved within them) and the impact of war-fighting and deployment upon recruitment. Nevertheless, it appears that contamination of training conditions was unlikely, since we did not see differences in skill acquisition across bases although cross-fertilization was not controlled.

An additional limitation of the current study includes the generalizability of our findings. The inclusion of only AF behavioral health specialists in the sample makes extrapolation to non-military samples uncertain. Compared to their civilian counterparts in training studies (Miller et al., 2004), AF behavioral health providers have less than half the years of general counseling experience and only two-thirds as much experience with substance abuse clients. This is probably a reflection of the shifting duty assignments of AF substance abuse treatment providers, who can be moved into client contact with very little formal training or experience. On the other hand, McLellan et al. (2003) note that substance abuse treatment programs typically have comparable provider turn over, meaning the high mobility of our participants might be analogous to their civilian counterparts. Further, our sample of military providers was substantially more ethnically diverse than that found in other published dissemination studies for MI, and therefore more likely to represent a civilian context. Finally, like their civilian counterparts, AF providers in this study were treating substance abuse clients largely mandated into treatment for alcohol-related infractions as opposed to seeking services voluntarily. Given the nature of their treatment population, lower levels of experience and limited skills at the onset of training, this sample may represent a fairly realistic estimate of the challenges in training complex empirically-based methods in public treatment agencies.

Taken as a whole, these data support the notion that training in MI can produce moderate improvements in counselor skills, even within a relatively inexperienced group with lower levels of basic counseling skills. Our data indicate that training enrichments of the kind that have been shown to increase performance in more skilled and experienced providers (Miller et al., 2004) may not be sufficient to do so with a less skilled group, particularly when clinicians do not access them fully. Workshop training is likely to be a necessary but not sufficient condition to ensure competence in MI among treatment providers in front line substance abuse treatment settings.

### Acknowledgements

This research was supported by Department of Defense DAMD 17-01-1-0681. We wish to thank Karin Cedro for her exceptional dedication to the overall quality of the research conducted in this project.

### References

- Baer, J. S., Rosengren, D. B., Dunn, C. W., Wells, E. A., Ogle, R. L. and Hartzler, B. (2004). An evaluation of workshop training in motivational interviewing for addiction and mental health clinicians. *Drug and Alcohol Dependence*, 73, 99–106.
- Carroll, K. M., Ball, S. A., Nich, C., Martino, S., Frankforter, T. L., Farentinos, C., Kunkel, L. E., Mikulich-Gilbertson, S. K., Morgenstern, J., Obert, J. L., Polcin, D., Snead, D. and Woody, G. E. (2006). Motivational interviewing to improve treatment engagement and outcome in

- individuals seeking treatment for substance abuse: a multisite effectiveness study. *Drug and Alcohol Dependence*, 81, 301–312.
- Carroll, K. M. and Rounsaville, B. J.** (2003). Bridging the gap: a hybrid model to link efficacy and effectiveness research in substance abuse treatment. *Psychiatric Services*, 54, 333–339.
- Cicchetti, D. V.** (1994). Guidelines, criteria and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6, 284–290.
- Hettema, J., Steele, J. and Miller, W. R.** (2005). Motivational interviewing. *Annual Review of Clinical Psychology*, 1, 91–111.
- McLellan, A. T., Carise, D. and Kleber, H. D.** (2003). Can the national addiction treatment infrastructure support the public's demand for quality care? *Journal of Substance Abuse Treatment*, 25, 117–121.
- Miller, W. R. and Mount, K. A.** (2001). A small study of training in motivational interviewing: does one workshop change clinician and client behavior? *Behavioural and Cognitive Psychotherapy*, 29, 457–471.
- Miller, W. R. and Rollnick, S.** (2002). *Motivational Interviewing: preparing people to change addictive behaviors* (2nd ed.). New York: Guilford Press.
- Miller, W. R., Rollnick, S. and Moyers, T. B.** (1998). *Motivational Interviewing: professional training videotape series*. University of New Mexico: Center on Alcoholism, Substance Abuse and Addictions.
- Miller, W. R., Sorenson, J. L., Selzer, J. A. and Brigham, G. S.** (2006). Disseminating evidence-based practices in substance abuse treatment: a review with suggestions. *Journal of Substance Abuse Treatment*, 31, 25–39.
- Miller, W. R., Yahne, C. E., Moyers, T. B., Martinez, J. and Pirritano, M.** (2004). A randomized trial of methods to help clinicians learn motivational interviewing. *Journal of Consulting and Clinical Psychology*, 72, 1050–1062.
- Moyers, T. B., Martin, T., Manuel, J. K., Hendrickson, S. M. L. and Miller, W. R.** (2005). Assessing competence in the use of motivational interviewing. *Journal of Substance Abuse Treatment*, 28, 19–26.
- Moyers, T. B., Martin, T., Manuel, J. K. and Miller, W. R.** (2003). The motivational interviewing treatment integrity (MITI) code. Version 2.0. Available <http://www.casaa.unm.edu/download/miti.pdf>. Retrieved October 25, 2006.
- Najavits, L. M., Crits-Christoph, P. and Dierberger, A.** (2000). Clinicians' impact on the quality of substance use disorder treatment. *Substance Use and Misuse*, 35, 2161–2190.
- Rubell, E. C., Sobell, L. C. and Miller, W. R.** (2000). Effectiveness of a workshop on motivational interviewing with alcohol abusers. *Behavior Therapist*, 23, 73–77.
- Schoener, E. P., Madeja, C. L., Henderson, M. J., Ondersma, S. J. and Janisse, J. J.** (2006). Effects of motivational interviewing training on mental health therapist behavior. *Drug and Alcohol Dependence*, 82, 269–275.
- Shafer, M. S., Rhode, R. and Chong, J.** (2004). Using distance education to promote the transfer of motivational interviewing skills among behavioral health professionals. *Journal of Substance Abuse Treatment*, 26, 141–148.
- Smith, J. L., Amrhein, P. C., Brooks, A. C., Carpenter, K. M., Levin, D., Schreiber, E. A., Travaglini, L. A. and Nunes, E. V.** (2007). Providing live supervision via teleconferencing improves acquisition of motivational interviewing skills after workshop attendance. *The American Journal of Drug and Alcohol Abuse*, 33, 163–168.
- Tappin, D. M., McKay, C., McIntyre, D., Gilmour, W. H., Cowan, S., Crawford, F., Currie, F. and Lumberson, M.** (2000). A practical instrument to document the process of motivational interviewing. *Behavioural and Cognitive Psychotherapy*, 28, 17–32.
- Walters, S. T., Matson, S. A., Baer, J. S. and Ziedonis, D. M.** (2005). Effectiveness of workshop training for psychosocial addiction treatments: a systematic review. *Journal of Substance Abuse Treatment*, 29, 283–293.