HHS Public Access

Author manuscript

J Subst Abuse Treat. Author manuscript; available in PMC 2016 October 01.

Published in final edited form as:

J Subst Abuse Treat. 2015 October ; 57: 96–101. doi:10.1016/j.jsat.2015.04.006.

Self-report after randomly assigned supervision does not predict ability to practice Motivational Interviewing

R. Morgan Wain,

Division of Substance Abuse, New York State Psychiatric Institute, 1051 Riverside Drive, Room 3732, Box 120, New York, NY 10032, Phone and fax: 650-641-2682

Bryan A. Kutner,

Division of Substance Abuse, New York State Psychiatric Institute, 1051 Riverside Drive, Room 3732, Box 120, New York, NY 10032, Phone: 415-596-9179, Fax: 650-641-2682

Jennifer L. Smith,

Department of Psychiatry, Columbia University, 1051 Riverside Drive, Room 3732, Box 120, New York, NY 10032, Phone: 212-543-4980, Fax: 212-543-6913

K. M. Carpenter,

Department of Psychiatry, Columbia University, New York State Psychiatric Institute, 1051 Riverside Drive, Room 3732, Box 120, New York, NY 10032, Phone: 212-923-3031

M-C. Hu,

Department of Psychiatry, Columbia University, 1051 Riverside Drive, Unit 20, New York, NY 10032, Phone: 212-304-7073

Paul C. Amrhein, and

Department of Psychiatry, Columbia University, New York State Psychiatric Institute, 1051 Riverside Drive #120, New York, NY 10032, Phone: 201-835-7276, Fax: 973-655-5121

E. V. Nunes[*]

Department of Psychiatry, Columbia University, New York State Psychiatric Institute, 1051 Riverside Drive, Unit 51, New York, NY 10032, Phone: 646-774-6122, Fax: 212-543-5326

R. Morgan Wain: morgan.wain@gmail.com; Bryan A. Kutner: bkutner@uw.edu; Jennifer L. Smith: smitjen@nyspi.columbia.edu; K. M. Carpenter: carpent@nyspi.columbia.edu; M-C. Hu: mh2236@columbia.edu; Paul C. Amrhein: pa2146@columbia.edu; E. V. Nunes: nunesed@nyspi.columbia.edu

Abstract

Correspondence to: R. Morgan Wain, morgan.wain@gmail.com.

[*]Disclosures

Dr. Nunes receives medication and/or placebo for research studies from Alkermes/Cephalon, Inc., Reckitt-Benckiser, and Duramed Pharmaceuticals, and he has received or will receive from Brainsway devices under investigation and reimbursement for travel expenses for investigators' meeting in connection with a Brainsway research study; he was paid an honorarium and received reimbursement for travel expenses for attendance at a Lilly Advisory Board Meeting in January 2012. HealthSim, LLC supplied web-based behavioral intervention for a research study in data analysis. No other authors have conflicts of interest to disclose.

The objective of this study was to investigate the relation between self-report and objective assessment of Motivational Interviewing (MI) skills following training and supervision. After an MI workshop, 96 clinicians from 26 community programs (age 21–68, 65% female, 40.8% Black, 29.6% Caucasian, 24.5% Hispanic, 2.0% Asian, 3.1% other) were randomized to supervision (tele-conferencing or tape-based), or workshop only. At four time points, trainees completed a self-report of MI skill, using items from the MI Understanding questionnaire (MIU), and were objectively assessed by raters using the Motivational Interviewing Treatment Integrity (MITI) system. Correlations were calculated between MIU and MITI scores. A generalized linear mixed model was tested on MIU scores, with MITI scores, supervision condition and time as independent variables. MIU scores increased from pre-workshop (*Mean* = 4.74, *SD* = 1.79) to post-workshop (*Mean* = 6.31, *SD* = 1.03) ($t = 8.69$, $p < .0001$). With supervision, scores continued to increase, from post-workshop to week 8 (*Mean* = 7.07, *SD* = 0.91, $t = 5.60$, $p < .0001$) and from week 8 to week 20 (*Mean* = 7.28, *SD* = 0.94, $t = 2.43$, $p = .02$). However, MIU scores did not significantly correlate with MITI scores, with or without supervision. Self- reported ability increased with supervision, but self-report was not an indicator of objectively measured skill. This suggests that training does not increase correspondence between self-report and objective assessment, so community treatment programs should not rely on clinician self- report to assess the need for ongoing training and supervision and it may be necessary to train clinicians to accurately assess their own skill.

Keywords

motivational interviewing; training; community-based supervision; dissemination; evidence-based practice

1. Introduction

Motivational interviewing (MI) is a well established evidence-based practice (e.g., Brown & Miller, 1993; Burke, Dunn, Atkins, & Phelps, 2004; Lundahl, Kunz, Brownell, Tollefson, & Burke, 2010; Miller & Rollnick, 2002; Rubak, Sandbaek, Lauritzen, & Christensen, 2005; Vasilaki, Hosier, & Cox, 2006; Wain et al., 2011) that requires direct observation and feedback to develop proficiency (de Roten, Zimmermann, Ortega, & Despland, 2013; Martino, Canning-Ball, Carroll, & Rounsaville, 2011b; Miller & Rose, 2009; Miller, Sorensen, Selzer, & Brigham, 2006; Miller, Yahne, Moyers, Martinez, & Pirritano, 2004; Schoener, Madeja, Henderson, Ondersma, & Janisse, 2006; Smith et al., 2012). However, if clinicians could rate their ability to practice MI accurately, this would reduce the need for time-intensive observer ratings and feedback, thus making dissemination more time- and cost-effective.

To date, only a few studies have examined the relation between clinicians' self-reported ability to practice MI and their actual ability as assessed objectively by raters. In an early and small training study ($N = 15$), Miller and Mount established that clinicians' perceived understanding about MI after workshop was not associated with actual skill as measured by objective assessment, and was frequently consistent with overconfidence (Miller & Mount, 2001). In a later randomized clinical trial evaluating different measures of feedback and coaching, self-reported understanding of MI among clinicians ($N = 140$) had little relation to

actual proficiency. At best, positive correlations between self-reported understanding and direct measures of ability were significant but modest (ranging from .169 to .329) with at least one inverse relationship ($r = -.244$), suggesting unclear relations between self-report and ability (Miller et al., 2004). Additional studies suggest that clinician self-report of fidelity to evidence-based practice tends to be higher than independent rater evaluations (Carroll, Nich, & Rounsaville, 1998; Decker & Martino, 2013; Martino, Ball, Nich, Frankforter, & Carroll, 2009; Miller, Yahne, & Tonigan, 2003).

Martino and colleagues advanced the field by conducting a study evaluating the correspondence of assessment, in which clinicians, their supervisors, and independent raters all rated clinicians' performance using the same instrument for assessing MI skill. The findings by Martino et al. (2009) were consistent with earlier research (Miller et al., 2004; Miller & Mount, 2001) but added that, relative to observers, *both* clinicians and their supervisors were more positive in their evaluations of the degree to which the intervention was present and skillfully delivered (i.e., adherence and competence) (Carroll et al., 2000). More recently, Decker and Martino (2013) evaluated the relation between community-based clinicians' reported confidence in their ability to practice MI and their objectively-assessed MI skill following three training conditions (self-study; expert-led workshop and supervision; or workshop and supervision by expert-trained trainers from within the community program). Regardless of condition, self-report did not comport with objectively assessed ability, with the exception that increased confidence was associated with slightly increased competence in advanced MI strategies (e.g., addressing ambivalence).

That clinicians and even their supervisors do not accurately evaluate MI skill poses a significant problem for efforts to train clinicians and highlights the importance of evaluating whether other instruments or approaches might better lend themselves to correspondence between self-report and observer ratings. The present report is based on a larger study (Smith et al., 2012) which used the Motivational Interviewing Treatment Integrity (MITI) 2.0 code (Moyers, Martin, Manuel, Hendrickson, & Miller, 2005) to evaluate clinicians' MI skill over the course of workshop training and supervision, and affords another opportunity to evaluate clinicians' self-assessment. The MITI is a simpler tool than the Motivational Interviewing Skill Code (MISC), which was used in several of the prior studies (e.g., Miller et al., 2004; Miller & Mount, 2001), but like the MISC, the MITI collects moment to moment counts of key clinician utterances like open questions and reflections, as well as global scores of proficiency. The present study also randomly assigned clinicians to either workshop training alone, or workshop training followed by five individual supervision sessions, including written and verbal feedback, over the following eight weeks, a relatively intensive supervision regimen. This affords the opportunity to examine whether clinicians' self-assessment of MI skill better corresponds to objective assessment after a substantial course of training and supervision.

The present report aims to replicate and expand on earlier investigations by examining the relation between trainee self-reported and objectively-assessed ability in a randomized clinical trial involving community-based clinicians receiving workshop training in MI followed by assignment to receive (or not receive) more extensive supervision.

1.1 Hypotheses

Our hypotheses are that 1) as previously found, self-reported MI ability will initially not be associated with objectively-assessed ability, but that 2) following longer-term MI supervision, self-reported ability will be associated with objectively-assessed ability. We expect that written and verbal feedback during supervision will increase clinicians' ability to identify their level of skill, as supervision teaches clinicians to discern between MI and other counseling styles. In the current milieu in which MI is a well-known evidence-based practice, the hope would be that self-report could be a useful proxy for objective assessment in community settings once a clinician becomes proficient in MI.

2. Method

This report is based on a secondary analysis of a parent study that examined effects of different supervision conditions on the development of MI skills (Smith et al., 2012). In brief, participants attended a two-day MI training workshop and were then randomized to post-workshop supervision conditions. Participants assigned to tele-conferencing supervision completed five weekly practice counseling sessions. Each practice session included a simulated clinical interaction with an actor portraying a standard patient. The session was simultaneously monitored by a supervisor, who provided real-time feedback using tele-conferencing technology, and who provided follow-up written feedback. Written feedback included graphical scores and a narrative that outlined areas of strength and improvements needed. Participants assigned to tape-based supervision completed five weekly audio taped practice counseling sessions with actors portraying standard patients. The audio tape was then sent to a supervisor who provided the same type of written feedback, as well as verbal feedback via telephone to follow up on the written feedback, to provide opportunities for role-play and to address any questions or concerns the trainee may have had. Thus, participants in both supervision groups completed five practice sessions and received written and verbal feedback five times. Participants in the workshop only training condition received no feedback following the workshop.

All participants, those receiving either workshop training and supervision or workshop only, were assessed four times over the course of the training study: prior to the two-day workshop (pre-workshop), within seven days following completion of the workshop (week 1 post-workshop), and at weeks 8 and 20 following the workshop. All four assessments were based on audio taped clinical sessions between the clinician participant and a client enrolled at their treatment clinic. Each audio taped session was used to assess the clinicians' MI skill using the MITI.

The study was approved by the Institutional Review Board of the New York State Psychiatric Institute, as well as by the Institutional Review Boards of every participating community treatment program. All participating practitioners and clients gave written informed consent.

2.1 Main Outcomes of Parent Study

By the end of the study period, participants in the tele-conferencing supervision condition demonstrated higher Spirit and Empathy scores than those in the workshop only condition;

participants' scores in the tape-based supervision condition fell between those of the other two. Those in the tele-conferencing condition used fewer MI Non-Adherent behaviors and more MI Adherent behaviors and exhibited greater Reflection to Question Ratios than participants in the workshop only condition. The one unexpected finding was that those who received tape-based supervision demonstrated higher Percent Complex Reflection scores when compared to those who received tele-conferencing supervision. In addition, at each assessment time point, clinicians completed the self-assessment Motivational Interviewing Understanding questionnaire (MIU) (Miller & Mount, 2001).

2.2 Practitioner-participants

Practitioners from 26 substance abuse community treatment programs affiliated with the Long Island and New York Nodes of the NIDA Clinical Trials Network were invited to participate. Potential participants had to be between 18 and 75 years of age, to provide counseling services directly to clients, and to be employed at least half time at their treatment facility prior to enrolling in the study. Practitioners could not participate if they had attended an MI workshop in the past three months, received training to be an MI trainer, had participated in a previous MI research trial, or planned to leave their job within the coming six months.

One hundred counselors were consented by study personnel; three participants who did not complete the workshop and one without MI skills data were excluded from this analysis ($N = 96$). Participants were randomized into one of three groups: tele-conferencing supervision ($n = 31$), tape-based supervision ($n = 32$), and workshop only ($n = 33$). The randomization was stratified by type of clinic site (methadone maintenance, outpatient drug-free, residential). The present study combined the two supervision groups, teleconferencing and tape-based, in order to compare supervision to workshop only, as the aim was not to compare types of supervision, but to compare those who received supervision to those who did not.

Of the 96 drug treatment counselors, 37 had a master's degree or higher, while 56 had a bachelor's degree or lower (3 missing data). Site type: 50 worked in methadone maintenance treatment programs, 32 in outpatient drug-free treatment programs, and 14 in residential treatment programs. Mean age was 43.5 years (range 21–68; 10 missing cases), 66% female, 25% Hispanic, 40.6% Black, 29.2% Caucasian, 2.1% Asian, and 3.1% other.

2.3 Supervisors

Five doctoral level psychologists who had attended a two-day Motivational Interviewing workshop and received training on scoring the MITI for supervision purposes provided supervision. Supervisors submitted audio-taped clinical interactions for objective assessment by an independent rater. Supervisors were required to score at competency level or greater before supervising participants. Competency levels are described in the MITI 2.0 (Moyers et al., 2005), and are outlined in section 2.4.2 below.

2.4 Measures

2.4.1 Perceived ability in Motivational Interviewing—Self-reported understanding of MI skills at each time point was collected using several items from the MIU, which draws from a series of questions regarding listening and MI proficiency (Miller & Mount, 2001). The items used for this assessment specifically ask about perceived MI competency:

> I understand the basic ideas and principles of motivational interviewing.
>
> I feel proficient and able to use motivational interviewing in my practice.
>
> I am already using motivational interviewing in my work.
>
> I am a skillful good listener in working with my clients.
>
> If a client is not initially motivated, I feel able to increase his or her motivation.

Participants rated these items on a scale from *0* (*not at all*) to *9* (*very much*), thus higher scores indicated self-perception of greater MI skill. We conducted a principal component analysis with orthogonal rotation (resulting in uncorrected components) and confirmed only one component with eigenvalues > 1.0, which accounted for 59% of the total variance. The factor loading of this component was higher on the first three items (loading = .90) than the last two items (loading = 0.44–0.55). Thus a single MIU total score was calculated for each participant at each time point. The coefficient Cronbach's Alpha of this 5-item scale was 0.80.

2.4.2 Assessed ability in Motivational Interviewing—MI competency was assessed using the MITI Version 2.0 (Moyers et al., 2005). The MITI was empirically derived from a standard patient analysis of the lengthier MISC instrument to capture the major elements of MI practice. The MITI enables the rater, with a single "pass" of a tape-recorded session, to code a number of variables: two global ratings ranging on a scale from *1* to *7* (*MI Spirit* and *Empathy*) and frequency counts of therapist behaviors, including counts of *Open Questions (OQ)*, *Closed Questions (CQ)*, *Simple Reflections (SR)*, *Complex Reflections (CR)*, *MI Adherent behaviors (MIA)* (e.g., affirmations, reinforcing choice), and *MI Non-Adherent behaviors (MINA)* (e.g., confronting, advising). These behavior counts are used to generate several summary scores, which are described in the MITI 2.0 (Moyers, et al., 2005): 1) *Percent MIA* (MIA/(MIA + MINA) × 100); 2) *Percent Open Questions* (OQ/(OQ+CQ) × 100); 3) *Percent Complex Reflections* (CR/(CR+SR) × 100); and 4) *Reflection to Question Ratio* ((SR+CR)/(CQ+OQ) × 100). (Note that Open Questions and Simple and Complex Reflections are considered to be hallmarks of MI practice; Closed Questions are counted in order to determine the Percent Open Questions but are not, in and of themselves, considered to be MI behaviors.)

To achieve "competence" on the MITI 2.0, an MI practitioner must score ≥ 6 on MI Spirit and Empathy, have a Reflection to Question ratio ≥ 2, Percent Open Questions ≥ 70%, Percent Complex Reflections ≥ 50%, and Percent MIA = 100% (i.e., no MI Non-Adherent behaviors).

In addition to Percent MIA, we included an MINA behavior count in our analysis because we observed a large number of MINAs (and very few MIAs) in the population of counselors

that participated in this training program. The philosophy of many of the treatment programs was based upon a more traditional approach, and included the use of direction, instruction, advice and corrective feedback. These sessions would yield Percent MIA scores of zero, regardless of how many MINAs were present, so we found it useful to report both Percent MIA and number of MINAs.

2.5 Assessment Ratings

Five supervisors served as independent raters of MITI scores at the four assessment time points. Two of the supervisors were trained by the MITI training programs at the University of New Mexico; three completed the "Training New Trainers" course of the Motivational Interviewing Network of Trainers (MINT). Sessions were randomly assigned with the qualifier that raters did not rate the participants they had supervised during the trial. Raters were blind to the group assignment and time point of the session being rated. Ten percent of assessment sessions were rated by two different raters for reliability estimates. Coders discussed coding questions during weekly staff meetings. As reported in the primary outcomes paper (Smith, et al., 2012), Intra-Class Correlation coefficients were calculated for a fixed set of raters (Shrout & Fleiss, 1979) for the MITI global scores and behavioral counts and ranged from fair to good, according to Cicchetti's guidelines (Cicchetti, 1994): MI Spirit 0.60; Empathy 0.50; Open Questions 0.74; Simple Reflections 0.40; Complex Reflections 0.64; MI Adherent (MIA) 0.47; and MI Non-Adherent (MINA) 0.72.

2.6 Analysis

The mean MIU scores were calculated at four time-points and paired t-tests were used to test for mean changes over time. To test our two hypotheses that self-report of MI ability does not correlate with objectively-assessed MI skill initially but improves with supervision, we calculated Pearson correlations to assess the relation between total MIU scores and MITI indices: Spirit and Empathy, Percent MI Adherent statements, Percent Open Questions, and Percent Complex Reflections. Spearman correlations were used to assess the relation between total MIU scores and number of MI Non-Adherent statements and Reflection to Question Ratio, due to non-normally distributed data for MI Non-Adherent statements and Reflection to Question Ratio. Correlations were calculated among the total sample and within each training group at each of the four assessment points. Bonferroni correction was used and the hypotheses were tested at significance level $p < .007$.

To further evaluate the association between MIU and MITI scores by supervision, we tested a generalized linear mixed model on MIU scores at four time points (pre-workshop, post-workshop, week 8 post-workshop and week 20 post-workshop) including independent variables: MITI scores, supervision condition (tele-conferencing supervision and tape-based supervision were combined for analysis into one condition called Supervision vs. Workshop only), time, and interaction between time and supervision. Seven indicators of MI skill (Spirit, Empathy, number of MI Non-Adherent statements, Percent MI Adherent statements, Percent Open Questions, Percent Complex Reflections, and Reflection to Question Ratio) were included separately in the model on MIU scores. Because MI Non-Adherent statements and Reflection to Question Ratio were not normally distributed, the scores were logarithm transformed. The three-way and two-way interactions between MITI scores, supervision

condition and time were tested and remained in the model if significant at $p < .05$. The GLM methodology was chosen to handle within-subject correlations (Diggle, Liang, & Zeger, 2004). PROC GENMOD in SAS 9.3 was used to conduct the analysis.

3. Results

Table 1 shows the means and standard deviations of self-reported MI skill (MIU score) and objectively-assessed MI skill (MITI scores), and correlations between the two at four time-points among the total sample and by two supervision conditions. Overall, MIU scores increased from pre-workshop (*Mean* = 4.72, *SD* = 1.79, indicating a moderate level of self-reported skill), to post-workshop (*Mean* = 6.31, *SD* = 1.03, indicating a higher level of self-reported skill) ($t = 8.69$, $p < .0001$). MIU scores continued to increase after supervision, but to a lesser degree by week 8 (*Mean* = 7.07, *SD* = 0.91, $t = 5.60$, $p < .0001$), and by week 20 (*Mean* = 7.28, *SD* = 0.94, $t = 2.43$, $p = .02$). In each linear MIU model, the interaction between supervision condition by time was not significant, indicating no difference in mean MIU score between the two supervision conditions. Consistent with our first hypothesis, MIU scores did not significantly correlate with MITI scores before supervision, across all MITI indices. Contrary to our second hypothesis, association between clinicians' self-reported ability to practice MI and objectively-assessed ability was not significantly different in those who received supervision compared to those who did not. According to correlation coefficients as described by Cohen (Cohen, 1992), most of the correlations were small ($r = +/- 0.1$) with a couple approaching large ($r = +/- 0.5$), across both supervision conditions (supervision and workshop only), and across time points. Some correlations were in the positive direction, others in the negative direction; none reached statistical significance (recall that a Bonferroni correction was used and statistical significance set at .007).

In the models of MIU score, there was no significant three-way interaction between MITI scores, supervision condition, and time, and no significant two-way interaction between MITI scores and supervision condition. MIU score was not associated with any of the MITI scores tested before or during the supervision period. The associations between MIU and MITI scores for counselors in the supervision group were neither significant nor stronger than the associations in the workshop only group, at any time-point tested. MIU scores were not associated with MI Spirit scores ($b = -0.01$, $se = 0.06$, $z = -0.18$, $p = 0.85$), Empathy scores ($b = -0.05$, $se = 0.07$, $z = -0.76$, $p = 0.45$), number of MI Non-Adherent statements ($b = 0.09$, $se = 0.11$, $z = 0.79$, $p = 0.43$), Percent MI Adherent statements ($b = -0.004$, $se = 0.002$, $z = -1.93$, $p = 0.06$), Percent Open Questions ($b = 0.002$, $se = 0.004$, $z = 0.42$, $p = 0.67$), Percent Complex Reflections ($b = -0.003$, $se = 0.003$, $z = 1.17$, $p = 0.24$), or Reflection to Question Ratio ($b = -0.23$, $se = 0.23$, $z = -0.97$, $p = 0.33$).

4. Discussion

This study replicates and expands upon earlier findings (Carroll et al., 1998; Decker & Martino, 2013; Martino et al., 2009; Miller et al., 2003; 2004; Miller & Mount, 2001) suggesting that MI practitioners' self-reported MI ability is a poor indicator of their actual ability. Participants' MIU scores (their self-reported ability to use MI) did increase with

training; however we found no correlation between MIU scores and MITI scores (objectively-assessed ability). Contrary to our expectation, even after five supervision sessions including written and verbal feedback, we found no evidence of an association between self-reported and objectively-assessed skill. Notably, many participants did not reach proficiency (the level below competency) in MI as measured by the MITI, prompting a (currently ongoing) follow-up study that includes an additional four supervision sessions over an additional eight weeks using clinic clients, following the five actor sessions. Thus we do not yet have the opportunity to look at the accuracy of self-report in more highly skilled community based trainees.

Even in the present environment, in which MI is more well-known than it was in the days of Miller and Mount's 2001 study, and in which MI is now a recommended evidence-based practice, our data are consistent with previous findings that clinicians' self-assessed skill is not a reliable measure of their objectively-assessed skill (Carroll et al., 1998; Decker & Martino, 2013; Martino et al., 2009; Miller et al., 2003; 2004; Miller & Mount, 2001). Since self-reported skill is evidently not a reliable proxy for objectively-assessed skill, community-based settings cannot rely on self-report as an indicator of MI competency and should employ observer-ratings to monitor and improve fidelity to MI practice.

We hypothesized that as clinicians received additional supervision and became more familiar with MI, their self-reported ability would more accurately reflect their actual ability. This was not the case. The supervision in this study emphasized feedback but did not train clinicians to assess MI skill by listening to or rating their own sessions (as recommended by experts in MI training, see (Miller & Rollnick, 2012; Rosengren, 2009)). The training package Motivational Interviewing Assessment: Supervisory Tools for Enhancing Proficiency (MIA:STEP), developed by the National Institutes of Drug Abuse (NIDA) and the Substance Abuse and Mental Health Services Administration (SAMHSA), likewise recommends training clinicians to use an objectively-assessed measure of MI skill (Martino et al., 2006); doing so may improve clinicians' ability to align their self-reported and actual abilities in MI.

Our findings lend support to the necessity of objective assessment to determine clinicians' fidelity to MI practice, and reaffirm the warning asserted by others, that self-report is not a reliable indicator of MI adherence or skill (Carroll et al., 1998; Decker & Martino, 2013; Martino et al., 2009; Miller et al., 2003; 2004; Miller & Mount, 2001). This has important implications for community treatment, where time and resources may constrain the amount of MI training and supervision provided. Nonetheless, in the absence of feedback from objective assessment, trainees and their supervisors are not likely to realize that trainees' self-perceived understanding of MI does not equate to proficiency in MI.

Importantly, competence in MI is tied to client improvement. Specific MI therapist behaviors predict client change talk (Amrhein, Miller, Yahne, Palmer, & Fulcher, 2003; Moyers, Martin, Houck, Christopher, & Tonigan, 2009; Pirlott, Kisbu-Sakarya, Defrancesco, Elliot, & Mackinnon, 2012), which correlates with positive health outcomes. Administrators of community treatment programs risk losing these positive outcomes if their clinicians believe they are offering MI yet are not actually equipped to practice MI with

fidelity. Likewise, administrators may come to believe that "MI does not work" if positive clinical outcomes are not achieved, even if these outcomes cannot reasonably be expected in the absence of actual ability to practice MI. Our findings point to the importance of objective assessment of MI ability in community settings, and suggest that using self-report as a proxy for ability may have a serious negative impact on public health.

4.1 Study Limitations and Directions for Future Research

This study was conducted with a modest sample size. It was based on an exploratory analysis and was not the main aim of the study from which the data were collected.

The MIU and MITI assess different aspects of clinical practice and neither tool may capture practice completely, which may bias tests of association between the two measures. The MIU inquires about general perception of one's ability (e.g., "*I understand the basic ideas and principles of motivational interviewing; I feel proficient and able to use motivational interviewing in my practice,*) but does not ask about specific skills (e.g., using Complex Reflections or Open Questions). In addition, self-report has limited accuracy generally (Vazire, 2008) and in other clinical practices (Davis, Mazmanian, Fordis, Van Harrison, Thorpe & Perrier, 2006) in addition to MI specifically (Carroll et al., 1998; Decker & Martino, 2013; Martino et al., 2009; Miller et al., 2003; 2004; Miller & Mount, 2001). In comparison, the MITI does assess for the use of specific MI skills (e.g., Complex Reflections and Open Questions), global MI qualities (MI Spirit and Empathy) and relies on observer ratings rather than self-report. The elements measured in the MITI were derived from factor analysis of the items in the MISC and the MITI has shown good sensitivity for detecting improvement in clinical practice following training (Moyers et al., 2005). The MITI, however, focuses exclusively on therapist competence and lacks the assessment of contextual elements and client language that the MISC might otherwise capture in a more holistic, and one could argue more accurate, assessment of clinical practice. The MITI also better assesses foundational MI skills than it does advanced skills, though that may not be much of a concern in the current application where many participants did not reach competence. In brief, the MIU questions measure self-reported ability to practice MI and a general sense of competence but do not inquire about mastery of specific skills (adherence). The MITI, on the other hand, better captures adherence than it does competence, but is still limited in its capacity to capture clinical practice as holistically as more extensive measures can. The difference in what these two instruments measure in comparison to one another, and in comparison to other more extensive tools, could therefore attenuate the expected association among the scores.

Due to the complexity of defining, training, and implementing MI e.g., (Miller, 2001; Rollnick, 2001), the field of practice has yet to establish an optimal assessment tool that evaluates effective practice both accurately and efficiently. Madson and Campbell (2006) compared a number of MI assessment measures and conclude, for example, that a clinician may reach competence as assessed by the MITI yet this does not assure skillful implementation or positive clinical outcomes. While the MIU and MITI do not comport in our study we cannot assume that the MITI is a comprehensive indicator of skill, limiting the conclusions that should be drawn based on their lack of comportment.

Taken together, the findings of this and related studies suggest that more work is needed to develop methods for training community-based clinicians in MI and related evidence-based approaches. Objective-assessment of clinicians' performance through independent raters, rather than through self-report would seem to be essential. A remaining question is whether clinicians could be trained to more accurately assess their own performance. The MITI, used by independent raters in the present study, provides detailed data regarding key MI consistent or inconsistent clinician behaviors, but clinicians cannot use a scale like the MITI while in the midst of a session. It is possible that training clinicians to use the MITI or a similar instrument retrospectively to rate their own or others' recorded sessions might be a fruitful training strategy (Martino et al., 2006; Miller & Mount, 2001; Miller & Rollnick, 2012; Rosengren, 2009). The Independent Tape Rater Scale (ITRS) used in studies by Martino and colleagues (Decker & Martino, 2013; Martino et al., 2009; 2011a) is another option that allows clinicians to retrospectively rate a session and elicits assessment of specific skills and MI strategies as well as global assessments. However, ITRS data are not as fine-grained in some respects as the counts yielded by the MITI.

An additional essential feature of MI is that the clinician must listen for and accurately assess the feedback received from the patient on a moment to moment basis during the interview (e.g., change talk or the absence thereof). Instruments like the DARN-C can be used to objectively rate client utterances with regard to change talk (Desires, Ability, Reasons, Need, and Commitment language), and the kind of client language evaluated with this instrument significantly predicts long term clinical outcomes (Amrhein et al., 2003). The sequence of interactions between counselor and patient is also important and a point at which accurate assessment could provide useful feedback on counselor adherence and competence (Amrhein et al., 2003; Moyers, Martin, Catley, Harris, & Ahluwalia, 2003). How to make such training for clinicians practical within the confines of a typical, busy community-based treatment program is another challenge. Emerging evidence suggests that client perception may be a practical and reliable alternative to direct observation of MI practice in community settings (Madson et al., 2013), where taping and observation are difficult (Schumacher, Madson, & Norquist, 2011). Group supervision with clinicians taking turns interviewing and rating others' interviews, eitherthrough role-play or monitoring of taped sessions, might be a sustainable framework (Forsberg, Ernst, Sundqvist, & Farbring, 2011; Mastroleo, Magill, Barnett, & Borsari, 2014; Schoener et al., 2006). In any case, future efforts should continue to examine and compare supervision and assessment methods to improve effective dissemination of evidence-based practices like Motivational Interviewing.

Acknowledgments

This work has been supported by the National Institute on Drug Abuse (NIDA) grants R01 DA016950 (Nunes), K24 DA022412 (Nunes) and K23 DA021850 (Carpenter).

References

Amrhein PC, Miller WR, Yahne CE, Palmer M, Fulcher L. Client commitment language during motivational interviewing predicts drug use outcomes. Journal of Consulting and Clinical Psychology. 2003; 71(5):862–878.10.1037/0022-006X.71.5.862 [PubMed: 14516235]

Brown JM, Miller WR. Impact of motivational interviewing on participation and outcome in residential alcoholism treatment. Psychology of Addictive Behaviors. 1993; 7(4):211–218.10.1037/0893-164X.7.4.211

Burke BL, Dunn CW, Atkins DC, Phelps JS. The Emerging Evidence Base for Motivational Interviewing: A Meta-Analytic and Qualitative Inquiry. Journal of Cognitive Psychotherapy. 2004; 18(4):309–322.

Carroll KM, Nich C, Sifry RL, Nuro KF, Frankforter TL, Ball SA, et al. A general system for evaluating therapist adherence and competence in psychotherapy research in the addictions. Drug and Alcohol Dependence. 2000; 57(3):225–238.10.1016/S0376-8716(99)00049-6 [PubMed: 10661673]

Carroll K, Nich C, Rounsaville B. Utility of therapist session checklists to monitor delivery of coping skills treatment for cocaine abusers. Psychotherapy Research. 1998; 8(3):307–320.

Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychological Assessment. 1994; 6(4):284–290.10.1037/1040-3590.6.4.284

Cohen J. A power primer. Psychological Bulletin. 1992; 112(1):155–159. http://psycnet.apa.org/journals/bul/112/1/155.html. 10.1037/0033-2909.112.1.155 [PubMed: 19565683]

Davis DA, Mazmanian PE, Fordis M, Van Harrison R, Thorpe KE, Perrier L. Accuracy of physician self-assessment compared with observed measures of competence: a systematic review. Jama. 2006; 296(9):1094–1102.10.1001/jama.296.9.1094 [PubMed: 16954489]

de Roten Y, Zimmermann G, Ortega D, Despland JN. Meta-analysis of the effects of MI training on clinicians' behavior. Journal of Substance Abuse Treatment. 2013; 45(2):155–162.10.1016/j.jsat.2013.02.006 [PubMed: 23537923]

Decker SE, Martino S. Unintended effects of training on clinicians' interest, confidence, and commitment in using motivational interviewing. Drug and Alcohol Dependence. 2013; 132(3):681–687.10.1016/j.drugalcdep.2013.04.022 [PubMed: 23684633]

Diggle, P.; Liang, K-Y.; Zeger, S. Analysis of Longitudinal Data. New York: Oxford University Press; 2004.

Forsberg LG, Ernst D, Sundqvist K, Farbring CÅ. Motivational Interviewing delivered by existing prison staff: a randomized controlled study of effectiveness on substance use after release. Substance Use & Misuse. 2011; 46(12):1477–1485.10.3109/10826084.2011.591880 [PubMed: 21756034]

Lundahl BW, Kunz C, Brownell C, Tollefson D, Burke BL. A Meta-Analysis of Motivational Interviewing: Twenty-Five Years of Empirical Studies. Research on Social Work Practice. 2010; 20(2):137–160.10.1177/1049731509347850

Madson MB, Campbell TC. Measures of fidelity in motivational enhancement: a systematic review. Journal of Substance Abuse Treatment. 2006; 31(1):67–73.10.1016/j.jsat.2006.03.010 [PubMed: 16814012]

Madson MB, Mohn RS, Zuckoff A, Schumacher JA, Kogan J, Hutchison S, et al. Measuring client perceptions of motivational interviewing: factor analysis of the Client Evaluation of Motivational Interviewing scale. Journal of Substance Abuse Treatment. 2013; 44(3):330–335.10.1016/j.jsat.2012.08.015 [PubMed: 22999814]

Martino, S.; Ball, SA.; Gallon, SL.; Hall, D.; Garcia, M.; Ceperich, S., et al. Motivational Interviewing Assessment: Supervisory Tools for Enhancing Proficiency. Salem, OR: Northwest Frontier Addiction Technology Transfer Center, Oregon Health and Science University; 2006. p. 1-246.

Martino S, Ball SA, Nich C, Canning-Ball M, Rounsaville BJ, Carroll KM. Teaching community program clinicians motivational interviewing using expert and train-the-trainer strategies. Addiction. 2011a; 106(2):428–441.10.1111/j.1360-0443.2010.03135.x [PubMed: 20925684]

Martino S, Ball S, Nich C, Frankforter TL, Carroll KM. Correspondence of motivational enhancement treatment integrity ratings among therapists, supervisors, and observers. Psychotherapy Research. 2009; 19(2):181–193.10.1080/10503300802688460 [PubMed: 19396649]

Martino S, Canning-Ball M, Carroll KM, Rounsaville BJ. A criterion-based stepwise approach for training counselors in motivational interviewing. Journal of Substance Abuse Treatment. 2011b; 40(4):357–365.10.1016/j.jsat.2010.12.004 [PubMed: 21316183]

Mastroleo NR, Magill M, Barnett NP, Borsari B. A pilot study of two supervision approaches for peer-led alcohol interventions with mandated college students. Journal of Studies on Alcohol and Drugs. 2014; 75(3):458–466. [PubMed: 24766758]

Miller WR. Comments on Dunn et al.'s 'The use of brief interventions adapted frommotivational interviewing across behavioral domains: a systematic review'. Addiction. 2001; 96(12):1769–1775.10.1080/09652140120089517 [PubMed: 11784469]

Miller WR, Mount KA. A small study of training in motivational interviewing: Does one workshop change clinician and client behavior? Behavioural and Cognitive Psychotherapy. 2001; 29(04): 457–471.

Miller, WR.; Rollnick, S. Motivational Interviewing. 2. New York: Guilford Press; 2002.

Miller, WR.; Rollnick, S. Motivational Interviewing. 3. New York: Guilford Press; 2012.

Miller WR, Rose GS. Toward a theory of motivational interviewing. American Psychologist. 2009; 64(6):527–537.10.1037/a0016830 [PubMed: 19739882]

Miller WR, Sorensen JL, Selzer JA, Brigham GS. Disseminating evidence-based practices in substance abuse treatment: A review with suggestions. Journal of Substance Abuse Treatment. 2006; 31(1): 25–39.10.1016/j.jsat.2006.03.005 [PubMed: 16814008]

Miller WR, Yahne CE, Tonigan JS. Motivational interviewing in drug abuse services: a randomized trial. Journal of Consulting and Clinical Psychology. 2003; 71(4):754–763. [PubMed: 12924680]

Miller WR, Yahne CE, Moyers TB, Martinez J, Pirritano M. A randomized trial of methods to help clinicians learn motivational interviewing. Journal of Consulting and Clinical Psychology. 2004; 72(6):1050–1062.10.1037/0022-006X.72.6.1050 [PubMed: 15612851]

Moyers TB, Martin T, Houck JM, Christopher PJ, Tonigan JS. From in-session behaviors to drinking outcomes: A causal chain for motivational interviewing. Journal of Consulting and Clinical Psychology. 2009; 77(6):1113–1124.10.1037/a0017189 [PubMed: 19968387]

Moyers TB, Martin T, Manuel JK, Hendrickson SML, Miller WR. Assessing competence in the use of motivational interviewing. Journal of Substance Abuse Treatment. 2005; 28(1):19–26.10.1016/j.jsat.2004.11.001 [PubMed: 15723728]

Moyers T, Martin T, Catley D, Harris KJ, Ahluwalia JS. Assessing the integrity of motivational interviewing interventions: Reliability of the motivational interviewing skills code. Behavioural and Cognitive Psychotherapy. 2003; 31(2):177–184.10.1017/S1352465803002054

Pirlott AG, Kisbu-Sakarya Y, Defrancesco CA, Elliot DL, Mackinnon DP. Mechanisms of motivational interviewing in health promotion: a Bayesian mediation analysis. The International Journal of Behavioral Nutrition and Physical Activity. 2012; 9(1):69.10.1186/1479-5868-9-69 [PubMed: 22681874]

Rollnick S. Comments on Dunn et al.'s 'The use of brief interventions adapted from motivational interviewing across behavioral domains: a systematic review'. Enthusiasm, quick fixes and premature controlled trials. Addiction. 2001; 96(12):1769–1770.10.1080/09652140120089517 [PubMed: 11784469]

Rosengren, DB. Building Motivational Interviewing Skills: A Practitioner Workbook. New York: Guilford Press; 2009.

Rubak S, Sandbaek A, Lauritzen T, Christensen B. Motivational interviewing: a systematic review and meta-analysis. Brit Journal of General Practice. 2005; 55(513):305–312.

Schoener EP, Madeja CL, Henderson MJ, Ondersma SJ, Janisse JJ. Effects of motivational interviewing training on mental health therapist behavior. Drug and Alcohol Dependence. 2006; 82(3):269–275.10.1016/j.drugalcdep.2005.10.003 [PubMed: 16289396]

Schumacher JA, Madson MB, Norquist GS. Clinical Forum-Using Telehealth Technology to Enhance Motivational Interviewing Training for Rural Substance Abuse Treatment Providers: A Services Improvement Project. Behavior Therapist. 2011; 34(4):64.

Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychological Bulletin. 1979; 86(2):420–428. [PubMed: 18839484]

Smith JL, Carpenter KM, Amrhein PC, Brooks AC, Levin D, Schreiber EA, et al. Training substance abuse clinicians in motivational interviewing using live supervision via teleconferencing. Journal of Consulting and Clinical Psychology. 2012; 80(3):450–464.10.1037/a0028176 [PubMed: 22506795]

Vasilaki EI, Hosier SG, Cox WM. The efficacy of motivational interviewing as a brief intervention for excessive drinking: a meta-analytic review. Alcohol and Alcoholism (Oxford, Oxfordshire). 2006; 41(3):328–335.10.1093/alcalc/agl016

Vazire S, Mehl MR. Knowing me, knowing you: the accuracy and unique predictive validity of self-ratings and other-ratings of daily behavior. Journal of Personality and Social Psychology. 2008; 95(5):1202–1216.10.1037/a0013314 [PubMed: 18954202]

Wain RM, Wilbourne PL, Harris KW, Pierson H, Teleki J, Burling TA, Lovett S. Motivational interview improves treatment entry in homeless veterans. Drug and Alcohol Dependence. 2011; 115(1–2):113–119.10.1016/j.drugalcdep.2010.11.006 [PubMed: 21145181]

Table 1

Mean and Correlations Between MIU Scores and Seven MITI Indices at pre-workshop, and weeks 1, 8 and 20 post-workshop (*N*=96)

	Pre-workshop		Week 1 post-workshop		Week 8 post-workshop		Week 20 post-workshop	
Total Sample	**Mean (SD)**	**r (p-value)**	**Mean (SD)**	**r (p-value)**	**Mean (SD)**	**r (p-value)**	**Mean (SD)**	**r (p-value)**
MIU scores	4.72 (1.79)		6.31 (1.03)		7.07 (0.91)		7.28 (0.94)	
Spirit	4.11 (1.04)	−.08 (.48)	4.77 (1.20)	.00 (.99)	4.79 (1.09)	.02 (.83)	4.89 (1.12)	.02 (.99)
Empathy	4.72 (1.79)	−.22 (.05)	4.78 (1.14)	.00 (.99)	4.81 (1.03)	.07 (.56)	4.85 (1.05)	−.02 (.88)
MI Non-Adherent	2.54 (3.72)	.03 (.82)	1.08 (2.54)	.22 (.05)	1.07 (2.17)	−.01 (.93)	1.06 (2.33)	−.06 (.61)
% MI Adherent	48.42 (36.47)	−.21 (.09)	70.34 (39.29)	−.08 (.53)	71.40 (39.08)	−.06 (.62)	77.13 (33.42)	.03 (.81)
% Open Questions	44.28 (20.24)	−.04 (.72)	53.57 (22.41)	.12 (.26)	58.47 (21.03)	.04 (.73)	51.94 (18.40)	.11 (.33)
% Complex Reflection	36.99 (26.01)	−.15 (.20)	45.45 (22.72)	−.17 (.12)	44.72 (22.80)	.02 (.86)	44.81 (25.64)	−.04 (.76)
Reflection to Question Ratio	0.47 (0.36)	−.06 (.58)	0.77 (0.75)	−.13 (.24)	0.83 (0.70)	−.19 (.10)	0.74 (0.77)	.00 (.99)
N	71–88	66–79	72–91	65–85	73–89	62–77	72–89	66–82
Workshop only group								
MIU scores	5.04 (1.76)		6.36 (1.10)		6.91 (1.00)		7.00 (1.13)	
Spirit	4.02 (0.75)	−.05 (.78)	4.88 (1.02)	−.21 (.24)	4.32 (0.87)	−.13 (.52)	4.63 (1.07)	−.04 (.85)
Empathy	4.22 (0.91)	−.43 (.02)	4.88 (0.96)	−.19 (.28)	4.42 (0.76)	−.07 (.74)	4.61 (1.02)	−.09 (.62)
MI Non-Adherent	1.19 (1.63)	.03 (.86)	0.61 (1.46)	.31 (.08)	1.32 (2.04)	−.04 (.86)	1.38 (2.55)	.08 (.66)
% MI Adherent	48.73 (39.27)	−.31 (.16)	73.26 (40.75)	−.12 (.57)	71.40 (39.08)	−.06 (.62)	68.90 (36.56)	−.07 (.75)
% Open Questions	48.59 (22.87)	−.37 (.05)	52.58 (24.35)	.14 (.42)	56.53 (44.52)	.00 (.99)	49.57 (17.77)	.23 (.21)
% Complex Reflection	37.13 (27.29)	−.24 (.22)	45.91 (20.52)	−.45 (.008)	43.81 (21.34)	.32 (.11)	48.80 (25.02)	.12 (.54)
Reflection to Question Ratio	0.49 (0.36)	−.28 (.14)	0.74 (0.49)	−.13 (.46)	0.67 (0.54)	−.30 (.13)	0.62 (0.37)	.06 (.74)
N	23–32	22–29	24–33	24–33	25–31	21–27	26–30	24–32
Supervision group								
MIU scores	4.55 (1.81)		6.28 (1.00)		7.15 (0.85)		7.44 (0.77)	
Spirit	4.16 (1.18)	−.07 (.62)	4.72 (1.29)	.11 (.44)	5.04 (1.12)	.04 (.79)	5.04 (1.13)	−.02 (.87)
Empathy	4.35 (1.24)	−.14 (.34)	4.72 (1.23)	.10 (.50)	5.03 (1.09)	.08 (.59)	4.99 (1.05)	−.02 (.86)
MI Non-Adherent	3.33 (4.34)	.06 (.66)	1.34 (2.97)	.16 (.26)	0.93 (2.25)	.05 (.73)	0.88 (2.20)	.14 (.32)
% MI Adherent	48.26 (35.47)	−.15 (.33)	68.85 (38.74)	−.06 (.69)	79.15 (33.90)	−.15 (.36)	81.78 (30.95)	.08 (.61)
% Open Questions	41.78 (18.30)	.15 (.31)	54.13 (21.28)	.11 (.45)	60.94 (21.30)	.00 (.99)	53.28 (18.76)	−.02 (.89)

J Subst Abuse Treat. Author manuscript; available in PMC 2016 October 01.

	Pre-workshop		Week 1 post-workshop		Week 8 post-workshop		Week 20 post-workshop	
Total Sample	**Mean (SD)**	**r (p-value)**	**Mean (SD)**	**r (p-value)**	**Mean (SD)**	**r (p-value)**	**Mean (SD)**	**r (p-value)**
% Complex Reflection	36.92 (25.48)	−.09 (.55)	45.19 (24.05)	−.01 (.92)	45.20 (23.71)	−.14 (.34)	42.57 (25.92)	−.09 (.50)
Reflection to Question Ratio	0.46 (0.36)	.08 (.58)	0.78 (0.86)	−.10 (.49)	0.91 (0.76)	−.15 (.29)	0.80 (0.92)	−.08 (.59)
N	48–58	44–50	48–58	41–52	48–58	41–50	46–57	42–52

All p-value > 0.007 (The significance cut-off with Bonferroni correction)