Published in final edited form as:

J Subst Abuse Treat. 2016 June ; 65: 20–25. doi:10.1016/j.jsat.2015.07.012.

Motivational Interviewing Fidelity in a Community Corrections Setting: Treatment Initiation and Subsequent Drug Use

Stephanie A. Spohr, MA¹, Faye S. Taxman, PhD², Mayra Rodriguez, PhD, MPH¹, and Scott T. Walters, PhD¹

¹University of North Texas Health Science Center, School of Public Health, Department of Behavioral and Community Health, Fort Worth, TX

²George Mason University, Department of Criminology, Fairfax, VA

Abstract

Introduction—Although substance use is common among people in the U.S. criminal justice system, treatment initiation remains an ongoing problem. This study assessed the reliability and predictive validity of the Motivational Interviewing Treatment Integrity 3.1.1. (MITI) coding instrument in a community corrections sample.

Methods—We used data from 80 substance-using clients who were participating in a clinical trial of MI in a probation setting. We analyzed 124 MI counseling sessions using the MITI, a coding system for documenting MI fidelity. Bivariate associations and logistic regression modeling were used to determine if MI consistent behaviors predicted substance use or treatment initiation at a two-month follow-up.

Results—We found a high level of agreement between coders on behavioral utterance counts. Counselors met at least beginning proficiency on most MITI summary scores. Probationers who initiated treatment at two-month follow-up had significantly higher ratings of clinician empathy and MI spirit than clients who did not initiate treatment. Other MITI summary scores were not significantly different between clients who had initiated treatment and those who did not. MI spirit and empathy ratings were entered into a forward logistic regression in which MI spirit significantly predicted two-month treatment initiation ($\chi^2(1) = 4.10, p < .05, R^2 = .05$) but counselor empathy did not. MITI summary scores did not predict substance use at two-month follow-up.

Conclusions—Counselor MI consistent relational skills were an important predictor of client treatment initiation. Counselor behaviors such as empathy and MI spirit may be important for developing client rapport with people in a probation setting.

Correspondence and requests to: Stephanie A. Spohr, MA, School of Public Health, The University of North Texas Health Science Center, 3500 Camp Bowie Ste 717, Fort Worth, TX 76104; sas0502@live.unthsc.edu; Phone: 763-670-7629.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conflict of Interest

None reported.

Keywords

Motivational Interviewing; Criminal Justice; Substance Abuse; Intervention Fidelity

1. Introduction

Among the nearly six million people under community supervision in the U.S. (i.e., probation and parole), two thirds are estimated to be drug involved (Taxman, Perdoni, & Caudy, 2013). However, despite the enormous leverage of the criminal justice system, less than half of drug-involved clients participate in substance abuse treatment (Karberg, 2005). The criminal justice system ranges from police interactions, to incarcerated settings, to community corrections. Although different in format, these formats also share a number of common challenges, including the need for strategies to motivate offenders to make changes that will reduce the likelihood of future criminal behavior. One way to improve substance use and treatment initiation rates may be to incorporate evidence-based counseling strategies such as motivational interviewing (MI) into criminal justice interactions (Walters, Clark, Gingerich, & Meltzer, 2007).

MI is a conversational style that originated from alcohol intervention studies in the early 1980's (Miller & Rollnick, 2012). MI has been widely validated as a stand-alone treatment, as a precursor to more extensive treatment, or integrated with other components, such as tailored feedback (Hettema, Steele, & Miller, 2005). Both relational features (e.g., counselor empathy) as well as specific language utterances (e.g., counselor choice of questions and statements) are thought to account for this positive effect of MI. MI has a strong evidence base in areas such as substance use, diet and exercise, and HIV risk behaviors (Lundahl & Burke, 2009). The positive effects of MI are generally robust across gender, age, ethnicity, and problem severity (Hettema et al., 2005). Additionally, there is a growing body of evidence that MI can affect criminal justice outcomes. Although this is a fairly new area, a review by McMurran (2009) found preliminary evidence that MI can increase motivation and treatment retention among criminal justice offenders, in both incarcerated and community corrections settings.

The link between counselor delivery of MI and outcome has been most strongly explored in the area of substance abuse (Apodaca & Longabaugh, 2009). However, studies have generally focused on treatment-seeking, rather than mandated populations. In fact, a comprehensive review of MI mechanisms of change did not identify any studies of clients that were mandated to attend treatment (Apodaca & Longabaugh, 2009). Mandated clients may be significantly different from voluntary populations both in terms of the client profile and in terms of system-level expectations. For instance, while legal penalties clearly provide incentives for action (or more accurately, punishments for inaction), they can also ignore many of the intrinsic reasons people would make changes (e.g., friends, family, personal responsibility), in a sense forcing people to act for external reasons (Fry, 2007). This may create a situation where provider interactions are focused more on short-term compliance tasks, or where clients adopt superficial language to avoid legal consequences.

Several measures have been developed to assess counselor fidelity to MI (Madson & Campbell, 2006). The most commonly used measure to document MI fidelity is the Motivational Interviewing Treatment Integrity coding system (MITI; Moyers, Martin, Manuel, Miller, & Ernst, 2010). The MITI is an in-progress (i.e., continual revisions) coding manual that measures how well a counselor adheres to MI principles. We used the most recent version available at the time of coding (3.1.1). The MITI measures five global scores of counselor behavior (i.e., evocation, collaboration, empathy, autonomy-support, and direction) and five counselor behavior counts that consist of giving information, questions (open and closed), reflections (simple and complex), and other utterances that are thought to be consistent or inconsistent with MI (MI adherent and MI non-adherent behaviors).

Scores on key MITI indicators have been found to be associated with clinical outcome. For instance, in a study with marijuana users, McCambridge, Day, Thomas, and Strang (2011) found that MI spirit and percent of complex reflections were predictive of marijuana cessation at a three-month follow-up. Other aspects of MI fidelity (i.e., counselor empathy, reflection to question ratio, percent open questions, and percent MI adherent) were not related to marijuana cessation. In a study with physically aggressive couples, Woodin, Sotskova, and O'Leary (2012) found that a greater reflection to question ratio and a greater percent open questions predicted reduced partner aggression at nine months. Greater counselor empathy was a marginally significant predictor, and other behaviors such as empathy, MI spirit, percent complex reflections and percent MI adherent utterances were not significantly associated with reduced partner aggression. Finally, in a study targeting people living with HIV, Seng & Lovejoy (2013) found that MI style and technique (determined via a principle component analysis of individual MITI indicators) was associated with fewer unprotected sex acts at six months. This overall pattern of findings suggests that some key MITI indicators can predict clinical outcome, however the indicators vary by study and the targeted behavior.

An important limitation of these studies is that they have drawn from community, rather than offender samples, specifically. In fact, a comprehensive review of MI mechanisms of change did not identify any studies that specifically focused on clients who were mandated to attend treatment (Apodaca & Longabaugh, 2009). One exception to this is work conducted in the Sweden prison system that found the MITI to be a reliable and valid measure of MI fidelity (Forsberg, Berman, Kallmen, Hermansson, & Helgason, 2008; Forsberg, Ernst, Sundqvist, & Farbring, 2011). Their research has indicated the MITI can reliably predict counselor training in MI (i.e., ongoing training with feedback versus service as usual). However, these studies also suggested that MI counselors had some difficulty meeting recommended MITI thresholds for beginning proficiency and competency with prisoners.

The current study examined the relationship between MI fidelity and clinical outcome in a group of probation clients. To our knowledge, this is the first study to explore the reliability and predictive validity of the MITI in a community corrections setting. The study had three aims: 1) Determine whether coders could reliably code MI interactions in a community corrections setting using the MITI; 2) Determine whether counselors could maintain fidelity to MI benchmarks in a community corrections setting, and 3) Examine the impact of counselors' MI consistent behaviors on subsequent substance abuse and treatment initiation

rates. We hypothesized that, consistent with the literature in other substance abusing and criminal areas, counselor behaviors that were consistent with the tenets of motivational interviewing (i.e., greater empathy and MI spirit, more complex than simple reflections, and more MI adherent behaviors) would predict reductions in substance use and increased treatment initiation.

2. Methods

Motivational Assessment Program to Initiate Treatment (MAPIT) was a multisite clinical trial comparing in person vs. computer approaches for increasing motivation to make changes in substance use, treatment initiation, and other behaviors related to probation success. The clinical trial is described more fully in Taxman et al. (in press). Briefly, probation clients at two sites (Dallas, TX and Baltimore City, MD) were stratified by criminal justice risk (low/moderate versus high risk) and randomized to: 1) A 2-session MI intervention (described here); 2) A 2-session motivational computer program, or 3) Supervision as usual. To be eligible, clients must have been at least 18 years old, within 30 days of their probation sentence date, and reported drug use or heavy alcohol use in the past 90 days. Participants were stratified at randomization by criminal justice involvement, a key predictor of criminal justice outcome (Monahan & Skeem, 2014). Clients completed a baseline interview (after which they received the first intervention visit), and follow-up visits at two and six months. This paper uses responses from the two-month follow-up to measure the proximal impact of MI on substance use and treatment initiation.

Participants receiving MI (N = 80) were 60% Black/African American, 20% Caucasian, 2.5% Native American, and 17.5% other or multiracial. Participants were mostly male (65%). Ages ranged from 18 to 57 years old (M = 34.9, SD = 12.0). Thirty-five percent of probationers were considered to be low risk for having further arrests (recidivism), 30% were moderate risk, and 35% were high risk. Twelve participants only received one MI session and six participants were lost to follow-up.

Briefly, the MI condition consisted of two, 45-minute sessions that incorporated a personalized feedback report (Walters, Ressler, Douglas, & Taxman, September, 2013). Session 1 (completed after the baseline interview) was intended to increase motivation to make changes in substance use, treatment initiation, and other early behaviors related to probation success. The Session 1 feedback report summarized different factors related to probation outcome, substance use, and included a section suggesting specific goals for the next month. Session 2 (completed approximately 30 days later) reviewed motivation and early progress, and offered additional tools to identify positive social support and short- and long-term goals. The Session 2 report was a worksheet (formatted similarly to Session 1) designed to help identify goals for the next month and assess social supports. Both feedback reports were intended to be completed collaboratively by the counselor and client. The client received a copy of their feedback report. Approximately half of MI sessions were completed in a private space in the probation office; the other half was completed at libraries, community centers, or other locations at the clients' request.

We used MI training procedures that were similar to previous clinical trials such as Project MATCH (Project MATCH Research Group, 1993). Counselors were trained and supervised by an experienced MI trainer (SW). There was one counselor located at each study site. The two counselors had master's degrees in psychology and social work, but had not been previously trained in MI. The counselors completed approximately 40 hours of training, including didactic information about MI spirit and technique, supervised role plays, and four practice sessions, prior to seeing clients. During the first three months of the study, counselors received weekly supervision with SW; thereafter, counselors received twice-monthly supervision. Tapes were coded per the instructions below.

2.1. Measures

2.1.1. Motivational Interviewing Treatment Integrity Scale—The MITI was originally derived from a factor analysis of the more comprehensive Motivational Interviewing Skill Code 1.0 measuring both client and counselor behaviors (MISC; Miller, Moyers, Ernst, & Amrhein, 2008). An exploratory factor analysis of the MISC 1.0 resulted in two global ratings of counselor behavior and five behavior utterance counts. The MITI 3.1.1 consists of five global scores assessing MI consistent behaviors (i.e., collaboration, autonomy support, evocation, direction, and empathy). These global scores are meant to capture the overall impression of each dimension on a 5-point Likert scale (1 = Low, 5 = High). The MITI also includes five behavior frequency counts: giving information, MI adherent, MI non-adherent, questions (i.e., open and closed), and reflections (i.e., simple and complex). The behavior counts are used to derive summary scores that serve as outcome measures for determining MI fidelity. The advantage of the MITI over more exhaustive coding systems (such as the MISC) is its reduced coding burden, particularly in non-research settings

The MITI 3.1.1 suggests thresholds for two levels of MI fidelity: *beginning proficiency* and *competency*. For instance, on the global scales, competency in MI is generally indicated by a score of at least 4.0 on a 5.0 scale. In terms of summary behavior counts, competency in MI is generally indicated by twice as many reflections as questions, 70% open questions (out of total questions), 50% complex reflections (out of total reflections), and 100% MI adherent utterances (out of the total MI adherent and non-adherent utterances).

The MITI has demonstrated good to excellent inter-rater reliability for the behavior count frequencies (McCambridge et al., 2011; Moyers, Martin, Manuel, Hendrickson, & Miller, 2005a; Pierson et al., 2007). There is mixed evidence regarding inter-rater reliability for the global scores. Some researchers have reported intraclass correlation coefficients for global scores lower than those of the behavior utterance counts (Forsberg, Kallmen, Hermansson, Berman, & Helgason, 2007; Moyers et al., 2005a; Pierson et al., 2007; Woodin, Sotskova, & O'Leary, 2012), which may be due to the restricted range for global measures (McCambridge et al., 2011; Moyers et al., 2005a; Pierson et al., 2007; Seng & Lovejoy, 2013)..

2.1.2. Timeline Follow-Back—The timeline followback (TLFB; Sobell & Sobell, 1996) is a widely utilized assessment tool that uses a calendar and other memory aids to assess

daily patterns of substance use and treatment attendance over a specified time period. The TLFB has often been used in lieu of biochemical measures when client invasiveness, cost, and time are an issue. The TLFB demonstrates good reliability and validity for assessing self-reported substance use (Fals-Stewart, O'Farrell, Freitas, McFarlin, & Rutigliano, 2000; Sobell, Brown, Leo, & Sobell, 1996). The TLFB demonstrates convergent and discriminant validity when compared to other measures and good overall agreement with urine analysis results in drug abusing clients (Fals-Stewart et al., 2000).

In this study, we used the TLFB to assess frequency of substance use at the two-month follow-up. Substances included: alcohol, marijuana, cocaine, opiates, hallucinogens, inhalants, barbiturates, amphetamines, and prescribed pain pills. Alcohol use was measured in standard drinks per day (i.e., 12 ounce beer, 5 ounce glass of wine, or 1.5 ounces of liquor), while illicit drug use was recorded as frequency of daily use.

2.2. MITI Coding Procedure

A total of 124 tapes were available for coding out of 172 possible sessions. Ten percent of sessions expired and were not completed per protocol, 14.5% of sessions were unrecorded, and 3.5% exhibited technical issues such as inaudible recording or tape shutting off prematurely. Slightly more than half (57%) were Session 1 tapes. Tapes were coded by four graduate students who had completed a semester-long class on MI skills and coding. MI session tapes were randomly assigned to coders. Tapes were coded in eight rounds. In each round, coders were assigned four unique tapes and one common tape. The common tape was randomly selected to assess for reliability and coders were blinded to the common tape. Tapes were generally coded in one pass, with coders reviewing difficult sections for accuracy. In between rounds, the coding team met to assess the commonly coded tapes, resolve discrepancies in coding, and prevent coder drift. Tapes were coded in their entirety to capture MI session characteristics over the course of a 45-minute session. Amrhein, Miller, Yahne, Palmer, and Fulcher (Amrhein, Miller, Yahne, Palmer, & Fulcher, 2003) indicate that important MI dynamics can occur towards the end of an MI session. Thus the entire session was coded rather than a random 10 – 20 minute segment. As the MI behavior outcome variables are ratios, the outcomes should not be affected by varying session lengths. Most session tapes lasted between 30 – 50 minutes.

2.3. MITI Predictor Variables

The MITI captures key MI fidelity measures through global scores and behavior frequency utterance counts. The MITI 3.1.1 manual suggests summary scores for determining MI fidelity. The first global summary measure encompasses the spirit of MI, a composite score of overall counselor collaboration, evocation, and autonomy-support. The second global summary measure consists of counselor empathy, capturing the clinician's interpersonal listening skills. A third global measure, direction, was omitted from analysis a priori because our intervention sessions followed a structured protocol with guided feedback sheets. Finally, four behavior count ratios include percent of MI adherent utterances out of MI adherent and nonadherent utterances, percent open questions out of total questions, percent complex reflections out of total reflections, and a reflection to question ratio.

2.4. Outcome Variables

TLFB data were used to classify clients into *responders* and *non-responders* regarding substance use and treatment initiation. With respect to *substance use*, clients were considered to be responders if they reported no illicit drug use and no heavy drinking episodes (i.e., 5 or more drinks for men or 4 or more drinks for women per day) at the two-month follow-up. Participants not meeting this criterion were coded as non-responders.

With respect to *treatment initiation*, clients were considered to be responders if they reported attending at least two self-help or substance abuse treatment sessions (i.e., anonymous self-help groups, group or individual counseling sessions, residential treatment, inpatient/outpatient treatment) at the two-month follow-up. Participants not meeting this treatment criterion were coded as non-responders. Two days of treatment attendance were used to indicate the direction of desired treatment attendance early in the probation process. Additionally, it helps to account for the single substance use assessment that is required of some probationers, which is viewed by some probationers as a “treatment” meeting.

2.5. Analysis Plan

Intraclass correlation coefficients (ICC) were calculated utilizing a two-way random effects model for consistency to assess inter-rater reliability. ICCs are a measure of inter-rater reliability commonly used to assess agreement ratings of continuous variables. Cicchetti (1994) indicates ICC scores below .4 are poor, between .4 and .59 are fair, between .6 and .74 are good, and above .75 is excellent. We used summary scores to assess the level of counselor MI proficiency by session on key MITI variables (i.e., global scores, percent MI-adherent, percent open questions, percent complex reflections, reflection to question ratio). For these portions of the analysis, we considered all available tapes.

T-tests were also conducted to assess the difference in MI consistent behaviors between positive responders (i.e., substance use, treatment initiation) and non-responders. For this portion of the analysis, we examined the combined effect of MI sessions by participants who had completed both sessions and the two-month follow-up interview. The MITI behavior counts were summed across both sessions while the global scores were averaged across sessions to keep the correct scaling. Following the bivariate association examination, significant MITI indicators were entered into a logistic regression to determine if MI consistent behaviors predicted substance use or treatment initiation at the two-month follow-up. For this portion of the analysis, we used all available tapes with two-month follow-up data.

3. Results

Inter-rater reliability ICC calculations are presented in Table A.1. The ICC analysis indicated an acceptable level of agreement between coders. Overall, 5/13 categories were in the excellent range, 4/13 were in the good range, and 4/13 were in the fair range. The MI non-adherent category displayed the lowest utterance agreement, indicating a greater difficulty identifying MI non-adherent utterances. Similar to other studies, we found poorer inter-rater agreement on the global assessment ratings compared to utterance counts due to

the restricted scoring range (Moyers et al., 2005a; Pierson et al., 2007; Seng & Lovejoy, 2013). However, in our study, the vast majority of the global ratings were a 4.0 or 5.0 on the 5-point scale, with 98% of ratings within ± 1 point value.

The mean MITI summary scores indicated good overall MI fidelity (see Table A.2). This sample included 71 MI Session 1 tapes and 53 MI Session 2 tapes. Mean clinician empathy was rated high and significantly improved from Session 1 to Session 2. Mean MI spirit was also rated high and significantly improved from Session 1 to Session 2. On average, the percent of MI adherent utterances, percent open questions, and percent complex reflections did not significantly improve from Session 1 to Session 2. Finally, the ratio of reflections to questions significantly *decreased* from Session 1 to Session 2. This is likely a result of the different structured format of Session 1 and 2; each session had specific topics and aims. Overall, counselors met MI competency levels for global empathy and MI spirit. Counselors met beginning proficiency levels for most other MITI summary scores, including percent MI adherent, percent open questions, and the reflection to question ratio. Only the percent of complex reflections was on average slightly below the beginning proficiency level.

Mean summary scores for combined MI sessions for probationers who completed both Session 1 and Session 2 are presented in Table A.3. This sample included a total of 80 tapes (40 clients). When we examined MITI scores in relation to client outcome, we found a significant difference in clinician empathy and MI spirit ratings between probationers who initiated treatment vs. those who did not. Overall, the MITI scores of sessions with clients who initiated treatment were higher in clinician empathy than clients who did not initiate treatment. MITI scores of sessions with clients who initiated treatment were higher in MI spirit ratings than clients who did not initiate treatment. MITI summary scores for clients who attended at least two treatment sessions at follow-up had higher ratings of clinician evocation, collaboration, and autonomy-support than clients who did not attend treatment. MITI summary scores were not significantly different between substance-using vs. non-using probationers, see Table A.4.

Clinician empathy and MI spirit ratings were entered into a forward Wald logistic regression prediction model for treatment initiation at two-months (see Table A.5). In the logistic regression, MI spirit ratings significantly predicted two-month treatment initiation ($\chi^2(1) = 4.10, p < .05, R^2 = .05$). For every one-unit increase in MI spirit rating, probationers were 2.82 times more likely to initiate early treatment efforts. Clinician empathy ratings did not significantly predict treatment initiation at two-month follow-up.

4. Discussion

This study assessed the reliability and predictive validity of the MITI in a sample of probation clients participating in a clinical trial. With regard to our first aim, we found the MITI was a reliable measure for rating MI fidelity in these interactions. The majority of ICC scores were in the acceptable range. Only MI non-adherent ratings did not reach agreement, indicating coder difficulty identifying and agreeing on MI non-adherent utterances. We also found a greater difficulty reaching agreement on counselor global measures due to the restricted scoring range. Our reliability correlations were similar to those reported in

previous studies in other populations (Forsberg et al., 2008; Forsberg et al., 2011; McCambridge et al., 2011; Moyers et al., 2005a; Pierson et al., 2007; Seng & Lovejoy, 2013).

With regard to our second aim, we found that study counselors can maintain fidelity to MI during a structured intervention in a community corrections setting. Counselors met MI competency levels for the MITI global scores and met beginning proficiency levels for most other MITI summary scores (i.e., percent MI adherent, percent open questions, reflection to question ratio). On average, counselors displayed some difficulty meeting the beginning proficiency level of percent of complex reflections. Forsberg et al. (2011) also found that counselors had difficulty meeting recommended beginning proficiency and competency levels with incarcerated substance abusing individuals. Importantly, the MITI does not take into account the context in which MI is delivered, and can underestimate counselor ability with particularly difficult clients. Moyers et al. (2005a) recommend assessing several samples of counselor functioning with difficult clients in order to draw firm conclusions regarding counselor competence. This study measured only the relationship between counselor fidelity to MI and client outcome. In the future, it will be important to assess client utterances using a measure like the MISC, in order to account for both counselor and client variables.

We found that counselor empathy and MI spirit ratings significantly improved from Session 1 to Session 2. This suggests that counselor rapport improved over time, which may have contributed to the observed positive outcomes. Also, the ratio of reflections to questions significantly decreased from Session 1 to Session 2. It is possible that this finding results from the slightly different objectives of the two sessions; Session 1 focused more on motivation for change with a heavier reliance on reflections, while Session 2 focused more on problem solving and goal setting relying more heavily on questions. Nevertheless, our overall MI fidelity was similar to that reported in previous literature (McCambridge et al., 2011; Moyers et al., 2005a; Pierson et al., 2007; Seng & Lovejoy, 2013).

Finally, with regard to our third aim, we found that MI adherent counselor relational behaviors predicted treatment initiation, but were unrelated to substance use in this sample. Contrary to McCambridge et al. (2011) who found that MITI scores predicted decreases in marijuana use in a community sample, we did not find that MITI scores predicted substance use in our setting. It is possible that difference resulted from our choice to use any drug use or heavy alcohol use, rather than volume of use, as the main outcome measure. It is possible that participants could have reduced their amount of substance use, but our study was not adequately powered to detect this effect. Additionally, the act of being under probation supervision can act as a deterrent in and of itself, which may have contributed to the null finding. We had expected that changes in treatment initiation would precede changes in drug use; thus, it is possible that two-month treatment initiation may lead to later changes in substance use as probationer's progress through treatment.

In this study, global scores of clinician empathy and MI spirit predicted early treatment efforts. Similar to our study, McCambridge et al. (2011) found MI spirit predicted marijuana cessation at a three-month outcome. Woodin, Sotskova, and O'Leary (2012) likewise

identified clinician empathy as an important predictor of reductions in partner aggression in women. We similarly conclude that a greater fidelity to MI spirit (i.e., collaboration, evocation, and autonomy-support) may be an important factor in motivating offenders to make early changes while on probation. In this study, other MITI summary scores such as percent MI adherent, percent open questions, percent complex reflections, and the reflection to question ratio did not predict treatment initiation at the two-month follow-up. Without neglecting the technical aspects of MI, this suggests that interventions delivered in this context may want to emphasize the client and counselor rapport (e.g., open questions, complex reflections). Because our results are preliminary, it will be necessary to further investigate these findings to determine the importance of various MI aspects for other outcomes including substance abuse.

4.1. Limitations

While this is the first study to report the reliability and predictive validity of the MITI with substance abusing offenders, there were some limitations to this study. First, because of our small sample size ($N = 40$), we may not have been powered to detect an effect of both global measures and behavior utterance counts. Likewise, due to the small sample size, we did not include covariates such as gender, age, criminal risk level, and baseline substance use in our analysis. Second, our analysis only included participants with two-month follow-up data. It is possible that participants who were lost to follow-up were different from those who were included in our sample. The small sample size in this study limits the validity, reliability, and generalizability of these findings. Third, this study only included two counselors (one per site), who received a relatively intensive training and supervision protocol. Although the intervention addressed a common probation goal, there are likely important differences between our intervention sessions and “typical” probation interactions, and thus we must be cautious about generalizing our findings beyond this relatively controlled setting. Fourth, our follow-up was limited to a relatively short-term, substance use/treatment outcome; we were not able to assess other distal outcomes such as compliance and recidivism. Finally, we observed lower reliability of global scores and MI non-adherent behavior counts while others have found these behaviors were correlated with outcome (Apodaca & Longabaugh, 2009). The low inter-rater reliability has the potential to undermine the validity of our findings or bias the estimate of association towards the null.

The MITI also has some inherent limitations for assessing MI. First, the MITI only assesses counselor functioning and does not take client variables into account; our findings do not provide additional explanation of why counselor variables might be related to client outcomes. Second, while the MITI is a good measure of counselor functioning, it may not be a good measure of purposeful and deliberate use of MI principles (Moyers et al., 2005a). In this study, Session 1 and Session 2 had slightly different objectives, and it is possible that the MITI may not be sensitive enough to account for their different objectives. Finally, some MITI measures have a limited range of values which limits conclusions we can make regarding counselor functioning and its predictive power (Moyers, Martin, Manuel, Hendrickson, & Miller, 2005b).

5. Conclusion

This study found good overall inter-rater agreement when using the MITI to assess counselor functioning in a sample of substance using probationers participating in a clinical trial. MI treatment fidelity was mostly excellent across counselors and sessions. This finding underscores the utility of the MITI as a tool for promoting evidence-based practice in the criminal justice system. It is important to note that our findings reflect the use of MI in a relatively structured setting in which counselors received continuous supervision with feedback. The importance of following up initial training with continuous feedback and coaching has been recognized in the broader literature (Schwalbe, Oh, & Zweben, 2014). There were interesting differences between Session 1 and 2 MITI summary scores, which suggested an overall improvement of clinician empathy and MI spirit from Session 1 to Session 2. This may suggest slightly different purposes of two kinds of structured MI interactions. Finally, we found that better MI spirit and empathy predicted two-month treatment initiation. This finding is important given that three million people in the criminal justice system are in need of substance abuse treatment. Identifying and reinforcing evidence-based practices such as MI can help reduce the health and safety costs associated with continued substance use.

Acknowledgements

This work was supported by a grant from the National Institute on Drug Abuse (R01 DA029010-01; Multiple PI: Walters/Taxman). The trial registration number is NCT01891656 (initiated on October 18, 2011). The authors gratefully acknowledge the contributions of our coding team: Brittany Marshall, Irene Reyes, Mayra Rodriguez and Teneshia Thurman.

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. [PubMed: 14516235]
- Apodaca TR, Longabaugh R. Mechanisms of change in motivational interviewing: A review and preliminary evaluation of the evidence. *Addiction*. 2009; 104(5):705–715. [PubMed: 19413785]
- Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normal and standardized assessment instruments in psychology. *Psychological Assessment*. 1994; 6:284–290.
- Fals-Stewart W, O'Farrell TJ, Freitas TT, McFarlin SK, Rutigliano P. The timeline followback reports of psychoactive substance use by drug-abusing patients: psychometric properties. *Journal of Consulting and Clinical Psychology*. 2000; 68(1):134–144. [PubMed: 10710848]
- Forsberg LG, Berman AH, Kallmen H, Hermansson U, Helgason AR. A test of the validity of the motivational interviewing treatment integrity code. *Cogn Behav Ther*. 2008; 37(3):183–191. doi: 794025509 [pii]10.1080/16506070802091171. [PubMed: 18608316]
- Forsberg LG, Ernst D, Sundqvist K, Farbring CA. Motivational Interviewing delivered by existing prison staff: a randomized controlled study of effectiveness on substance use after release. *Subst Use Misuse*. 2011; 46(12):1477–1485. [PubMed: 21756034]
- Fry R. Community corrections' core mission. *Corrections Today*. 2007; 69:14–14.
- Hettema J, Steele J, Miller WR. Motivational interviewing. *Annu Rev Clin Psychol*. 2005; 1:91–111. [PubMed: 17716083]
- Karberg, JC.; James, DJ. Substance dependence, abuse and treatment of jail inmates. Vol. 2002. Washington, DC: Dept of Justice Publication, NCJ 209588; 2005.

- Lundahl B, Burke BL. The effectiveness and applicability of motivational interviewing: a practice-friendly review of four meta-analyses. *Journal of Clinical Psychology*. 2009; 65(11):1232–1245. [PubMed: 19739205]
- Madson MB, Campbell TC. Measures of fidelity in motivational enhancement: A systematic review. *Journal of Substance Abuse Treatment*. 2006; 31(1):67–73. [PubMed: 16814012]
- McCambridge J, Day M, Thomas BA, Strang J. Fidelity to Motivational Interviewing and subsequent cannabis cessation among adolescents. *Addictive Behaviors*. 2011; 36(7):749–754. [PubMed: 21440994]
- McMurrin M. Motivational interviewing with offenders: A systematic review. *Legal & Criminological Psychology*. 2009; 14:83–100.
- Miller, WR.; Moyers, TB.; Ernst, D.; Amrhein, P. Manual for the motivational interviewing skill code (MISC). Version 2.1. 2008. <http://casaa.unm.edu/download/misc.pdf>
- Miller, WR.; Rollnick, S. Motivational interviewing: Helping people change. 3rd Ed.. New York: Guilford Press; 2012.
- Monahan J, Skeem JL. Risk redux: The resurgence of risk assessment in criminal sanctioning. *Federal Sentencing Reporter*. 2014; 26(3):158–166.
- Moyers TB, Martin T, Manuel JK, Hendrickson SM, Miller WR. Assessing competence in the use of motivational interviewing. *Journal of Substance Abuse Treatment*. 2005a; 28(1):19–26. [PubMed: 15723728]
- Moyers TB, Martin T, Manuel JK, Hendrickson SM, Miller WR. Assessing competence in the use of motivational interviewing. *Journal of Substance Abuse Treatment*. 2005b; 28(1):19–26. doi: S0740-5472(04)00123-0 [pii]10.1016/j.jsat.2004.11.001. [PubMed: 15723728]
- Moyers TB, Martin T, Manuel JK, Miller WR, Ernst D. Revised global scales: Motivational interviewing treatment integrity 3.1.1 (MITI 3.1.1). 2010 <http://casaa.unm.edu/download/miti3.pdf>.
- Pierson HM, Hayes SC, Gifford EV, Roget N, Padilla M, Bissett R, Fisher G. An examination of the Motivational Interviewing Treatment Integrity code. *Journal of Substance Abuse Treatment*. 2007; 32(1):11–17. [PubMed: 17175394]
- Project MATCH Research Group. Project MATCH: Rationale and methods for a multisite clinical trial matching patients to alcoholism treatment. *Alcoholism, Clinical and Experimental Research*. 1993; 17:1130–1145.
- Schwalbe CS, Oh HY, Zweben A. Sustaining motivational interviewing: a metaanalysis of training studies. *Addiction*. 2014; 109(8):1287–1294. [PubMed: 24661345]
- Seng KS, Lovejoy TI. Reliability and validity of a treatment fidelity instrument for motivational interviewing targeting sexual risk behaviors in people living with HIV/AIDS. *Journal of Clinical Psychology in Medical Settings*. 2013; 20:440–448. [PubMed: 23636311]
- Sobell LC, Brown J, Leo GI, Sobell MB. The reliability of the Alcohol Timeline Followback when administered by telephone and by computer. *Drug and Alcohol Dependence*. 1996; 42(1):49–54. [PubMed: 8889403]
- Sobell, LC.; Sobell, MB. Timeline Followback user's guide: A calendar method for assessing alcohol and drug use. Toronto, Ontario, Canada: Addiction Research Foundation; 1996.
- Taxman FS, Perdoni ML, Caudy M. The plight of providing appropriate substance abuse treatment services to offenders: Modeling the gaps in service delivery. *Victims and Offenders*. 2013; 8(1): 70–93.
- Taxman FS, Walters ST, Sloas LB, Lerch J, Rodriguez M. Counselor vs. computer: a randomized controlled trial protocol comparing motivational tools to improve probationer treatment initiation and reduce substance use. *Contemporary Clinical Trials*. (in press).
- Walters, ST.; Clark, MD.; Gingerich, R.; Meltzer, M. Motivating offenders to change: A guide for probation and parole. Washington DC: National Institute of Corrections, U.S. Department of Justice; 2007.
- Walters ST, Ressler E, Douglas L, Taxman FS. Motivational interviewing in criminal justice: A new approach to addressing treatment motivation and related behaviors. *Counselor: The Magazine for Addictions Professionals*. 2013 Sep.

Woodin EM, Sotskova A, O'Leary KD. Do motivational interviewing behaviors predict reductions in partner aggression for men and women? Behaviour Research and Therapy. 2012; 50(1):79–84. [PubMed: 22119133]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

- The MITI can be used reliably to assess MI fidelity in a criminal justice sample.
- Counselors can maintain fidelity to MI benchmarks with drug-involved offenders.
- Greater empathy and MI spirit predicted client treatment initiation at two months.
- Key MITI indicators did not predict abstinence at two months.

Table A.1

Inter-rater reliability of MITI indicators.

	ICC	<i>p</i>
Empathy	0.672	*
Autonomy-Support	0.688	*
Collaboration	0.505	NS
Evocation	0.405	NS
Giving Information	0.981	***
MI Adherent	0.694	*
MI Non-Adherent	0.463	NS
Closed Questions	0.794	**
Open Questions	0.946	***
Simple Reflections	0.864	***
Complex Reflections	0.741	**
Total Utterances	0.999	***

Note: *N* = 8;*
< 0.05;**
< 0.01;***
< 0.001;

NS – Not significant

Table A.2

Mean (SD) MITI summary scores by session.

	Beginning Proficiency Threshold	Competency Threshold	Session 1 (N = 71)	Session 2 (N = 53)	t	p
Empathy	3.5	4.0	4.29 (0.49)	4.52 (0.49)	3.13	**
MI Spirit	3.5	4.0	4.26 (0.41)	4.51 (0.39)	3.73	***
% MI Adherent	90%	100%	93.0 (15.2)	94.4 (18.1)	0.45	NS
% Open Questions	50%	70%	50.5 (13.6)	53.0 (9.40)	1.26	NS
% Complex Reflections	40%	50%	29.9 (9.40)	30.0 (12.1)	-0.47	NS
Reflections/Questions Ratio	1	2	1.50 (0.52)	1.15 (0.54)	-4.49	***

* < 0.05;

** < 0.01;

*** < 0.001;

NS – Not significant

Table A.3

Mean (SD) MITI summary scores by treatment attendance.

	Treatment (N = 13)	No Treatment (N = 27)	<i>t</i>	<i>p</i>
Combined Sessions				
Empathy	4.52 (0.33)	4.31 (0.29)	2.081	*
MI Spirit	4.50 (0.23)	4.29 (0.29)	2.349	*
% MI Adherent	90.5 (13.3)	89.3 (18.7)	0.198	NS
% Open Questions	52.8 (9.80)	51.2 (9.80)	0.469	NS
% Complex Reflections	30.5 (8.40)	29.5 (8.00)	0.384	NS
Reflections/Questions Ratio	1.41 (0.55)	1.26 (0.45)	0.950	NS

*
< 0.05;

NS – Not significant

Table A.4

Mean (SD) MITI summary scores by substance use outcome.

	Ceased Use (N = 24)	Continued Use (N = 16)	<i>t</i>	<i>p</i>
Combined Sessions				
Empathy	4.34 (0.31)	4.43 (0.32)	0.89	NS
MI Spirit	4.35 (0.31)	4.36 (0.24)	0.39	NS
% MI Adherent	91.5 (13.3)	87.2 (0.21)	-0.77	NS
% Open Questions	51.0 (9.6)	52.8 (10.1)	0.56	NS
% Complex Reflections	29.2 (7.5)	31.0 (9.1)	0.59	NS
Reflections/Questions Ratio	1.35 (0.53)	1.24 (0.42)	-0.69	NS

*
< 0.05;

NS – Not significant

Table A.5

Logistic regression model prediction of 2-month treatment initiation.

	B (SE)	OR (95% CI)	<i>p</i>
MI Spirit	1.04 (0.51)	2.82 (1.03, 7.70)	0.04
Constant	−5.22 (2.28)		