Journal of Offender Rehabilitation, 49:309–323, 2010
Copyright © Taylor & Francis Group, LLC
ISSN: 1050-9674 print/1540-8558 online
DOI: 10.1080/10509674.2010.489455

Routledge
Taylor & Francis Group

Motivational Interviewing as a Supervision Strategy in Probation: A Randomized Effectiveness Trial

SCOTT T. WALTERS, AMANDA M. VADER, NORMA NGUYEN, and T. ROBERT HARRIS

School of Public Health, University of Texas, Dallas, Texas, USA

JENNIFER EELLS

Dallas County Community Supervision and Corrections, Dallas, Texas, USA

Motivational interviewing (MI) has been recommended as a supervision style in probation. This project examined the effectiveness of an MI training curriculum on probation officer MI skill and subsequent probationer outcome. Twenty probation officers were randomized to receive MI training, or to a waiting list control, while an additional group of 10 officers served as a supervision-as-usual group; officer outcomes included questionnaires and standardized interactions at baseline, 2 months, and 6 months. A total of 380 probationers were assigned to officers during a 4-month period; offender outcomes included the probability of having a drug-positive urinalysis or an otherwise poor outcome after 6 months. The MI training program improved officer skill as measured by standardized interactions. However, after controlling for baseline characteristics, probationer outcome did not vary by training group, nor did officer MI competence predict outcome. Results are discussed in terms of the role of MI in the overall probation system.

KEYWORDS counseling, motivational interviewing, probation, rehabilitation

---

This project was funded through a cooperative agreement with the National Institute on Corrections (07C71GJS8, PI: Walters). The authors gratefully acknowledge the assistance of Malissa Cornett, Dale House, Vance Paulett, and Terri Thomas, as well as the 30 probation officers who donated their time to help implement this project.

Address correspondence to Scott T. Walters, School of Public Health, University of Texas, 5323 Harry Hines Blvd., V8.112, Dallas, TX 75390-9128, USA. E-mail: scott.walters@utsouthwestern.edu

INTRODUCTION

Motivational interviewing (MI) has recently been recommended as an evidence-based practice (EBP) in corrections to increase compliance with court mandates and encourage other healthy behavior changes, such as avoidance of drug and alcohol use (Alexander, VanBenschoten, & Walters, 2008; Clark, Walters, Gingrich, & Meltzer, 2006; Ginsburg, Mann, Rotgers, & Weekes, 2002; Wormith et al., 2007). The push toward EBP is supported by research suggesting that effective correctional programs share similar characteristics in terms of targeting offender risk, needs, and responsivity. For instance, a number of studies show that programs that match higher-risk offenders to more intensive services (e.g., risk) and address dynamic criminogenic factors (e.g., needs) can significantly reduce recidivism and other criminal behavior outcomes (Andrews & Bonta, 2003; Andrews et al., 1990). However, there has been much less research on the EBP principle of responsivity, which says that providers should use evidence-based communication strategies. In placing MI within an EBP framework, MI is best understood in terms of its consistency with the responsivity principle, since it suggests an interpersonal communication style.

MI is a "client centered, directive style of interacting with a person to help explore and resolve ambivalence about change" (Miller & Rollnick, 2002, p. 25). MI borrows from client-centered counseling in its emphasis on empathy, optimism, and respect for client choice (Rogers, 1961). MI also draws from self-perception theory, which says that a person becomes more or less committed to an action based on the verbal stance he or she takes (Bem, 1972). Thus, a person who argues in support of change is more likely to make that change, whereas a person who defends the status quo is more likely to persist in current behaviors. MI has strong research support in areas such as alcohol and drug use, smoking cessation, medication compliance, HIV risk behaviors, and diet/exercise (Lundahl, Tollefson, Kunz, Brownell, & Burke, 2010). Two meta-analyses (Hettema, Steele, & Miller, 2005; Rubak, Sandboek, Lauritzen, & Christensen, 2005) found that across behavior change areas, MI was significantly better than other approaches in three out of four published studies, and outperformed traditional advice-giving 80% of the time. Research from the addictions field suggests that MI may be particularly effective for those clients who are more oppositional or defiant, higher risk, or otherwise less ready for change (Project MATCH Research Group, 1998; Woodall, Delaney, Kunitz, Westerberg, & Zhao, 2007).

Several articles and book chapters have described the use of MI in criminal justice settings (Ginsburg et al., 2002; McMurran, 2002; Vivian-Byrne, 2004; Walters, Clark, Gingerich, & Meltzer, 2007), but there has been comparatively little outcome research on MI in corrections. In the largest review to date, McMurran (2009) examined 19 studies using MI among offender populations.

The review concluded that MI was generally effective at improving treatment engagement and retention, as well as improving substance abuse outcomes. However, the modal MI format in this review was as a discrete intervention in a specialized treatment setting; only one study examined MI as a probation supervision strategy. Harper and Hardy (2000) reported positive effects on probationers' problem recognition as a result of being assigned to an MI–trained officer, relative to probationers assigned to a non-MI trained officer. However, in this study random assignment was not mentioned, and the control group also showed some progress. Relatedly, two studies demonstrated improvement in probation officer MI skill following training. Miller and Mount (Miller & Mount, 2001) found modest improvements on some aspects of MI as judged by taped samples, while Hohman, Doran, and Koutsenok (2009) found that attendees significantly improved their knowledge of MI and reflective listening skills as judged on pen-and-paper measures. However, neither study measured changes in probationer outcome.

In addition to the relatively limited empirical literature, there are at least three practical reasons to believe that MI might be particularly effective at increasing compliance with court conditions in a community corrections setting. First, MI has strong empirical support in areas that may be relevant to community corrections, such as preparing clients to engage in alcohol and drug treatment programs (Baker et al., 2002; Daley, Salloum, Zuckoff, Kirisci, & Thase, 1998; Miller, Meyers, & Tonigan, 1999). Further, MI has been shown to be effective in other settings where provider-client interactions may be brief and multifocused, such as in medical consultations (Heather, Rollnick, Bell, & Richmond, 1996). Finally, large research trials such as Project MATCH (Project MATCH Research Group, 1998) have reported similar effects of MI on alcohol and drug use across offending and nonoffending clients. Thus, there has been an interest in translating aspects of MI into criminal justice settings, not only as a discrete counseling technique, but also in terms of an ongoing supervision strategy to encourage positive progress on probation conditions.

The present study evaluated MI as a supervision strategy in a large urban probation department. This is the first study of which we are aware to evaluate not only the effect of MI training on officer skill, but also the effect of MI training and skill on probationer outcome. This study had three aims: (a) determine whether a training program improved officer MI competence as judged by standardized role play interactions; (b) determine whether training group affected subsequent probationer compliance with probation conditions; and (c) determine whether officer MI competence was related to probationer compliance, independent of study group. The first two aims were designed to give us a picture of the effectiveness of the training program, while the third aim was designed to test whether any natural variation in officer MI skill would predict outcome. As an effectiveness trial, the project

was designed to mirror the real-world constraints that probation departments might face when implementing an MI training program without additional systemic changes. We recruited a group of officers with a regular (i.e., nonspecialized) caseload, trained a subset of these officers, and tracked new probationers assigned to them within a subsequent four-month period. We did not alter any aspect of client load, available resources, or other supervision policies. Participating officers were provided with training credit, but did not receive any compensation for their participation. Thus, we feel our results offer a good picture of the effects of a modest MI training program apart from any other systemic changes.

METHODS

Participants and Procedures

The study was conducted in four field offices of a large urban probation department. Twenty probation officers carrying a nonspecialized caseload who expressed interest in receiving MI training were randomized (stratified by office) to receive MI training or to a waiting list control group. Ten officers randomized to the MI trained (MI+) group completed an initial 24-hour MI training sequence with a member of the Motivational Interviewing Network of Trainers. Specifically, the training included an initial two-day workshop, a half-day "booster" training in the first month, and one or two monthly supervision meetings over six months. The training curriculum included lecture, demonstration, and practice in MI techniques that might be most useful in probation settings, as well as assistance with identifying specific behavioral targets. Officers randomized to the MI untrained (MI−) group continued to see clients as usual, but did not receive any MI training during the project period. A final group of 10 officers who were not interested in receiving MI training served as a supervision-as-usual (SAU) control group. The SAU group was included to test for any potential selection bias due to interest in participating in an MI training study. During the study period, officers continued to receive new cases assigned to them through the agency's normal zip code/alphabetical assignment process. This project was approved by the Human Subjects Board of the University of Texas Health Science Center at Houston; all officers provided voluntary informed consent prior to participating.

Measures (Probation Officers)

We measured officer MI skills with two questionnaires and a series of standardized interactions at three points: baseline, post (two months after baseline), and six months.

1. Officer reflective listening skill: The Officer Responses Questionnaire (ORQ; Walters, Alexander, & Vader, 2008) was included as a basic

measure of reflective listening capacity. The ORQ is a 5-item questionnaire in which an officer is asked to respond to a series of hypothetical probationer statements with a listening response. The ORQ is modified from the Helpful Responses Questionnaire, which has show adequate reliability in previous studies (Miller, Hedrick, & Orlofsky, 1991).

2. Officer beliefs: The Dual-Role Relationships Inventory-Revised (DRI-R; Skeem, Louden, Polaschek, & Camp, 2007) was used to measure officers' perceptions of their own helpfulness. The DRI-R, which consists of questions regarding perceived helpfulness, fairness, and an authoritative (not authoritarian) style, has been shown to be a reliable and valid indicator of the quality of the relationship between the probation officer and client. When measured through probationer (but not officer) self-report, DRI-R scores have predicted subsequent probation violations (Skeem et al., 2007).
3. Officer MI skill: As a proxy measure of MI skill, all officers completed two standardized role play interactions at three different time points (baseline, post, six months). In each case, officers were given a description of a medium-risk case and asked to conduct the interview as if it were a regular office visit. Scenarios (available from the authors upon request) described probationers who were ambivalent about making changes in substance abuse, anger management, or other behaviors closely related to probation; actors were typically probationers (from other field offices) who volunteered for the project to obtain community service hours. Probationer actors were not otherwise affiliated with the project and were not considered to be participants for human subjects purposes. Interactions were coded using the Motivational Interviewing Treatment Integrity (MITI) coding system (Moyers, Martin, Manuel, & Miller, 2003; Moyers, Martin, Manuel, Miller, & Ernst, 2007). The MITI has been widely used as a treatment integrity measure in MI intervention trials (Bennett, Roberts, Vaughan, Gibbins, & Rouse, 2007; Moyers, Martin, Manuel, Hendrickson, & Miller, 2005; Walters, Vader, Harris, Field, & Jouriles, 2009). Our coding procedures are described more fully elsewhere (Walters, Cornett, & Vader, in press). Briefly, each tape was coded by one of three randomly-assigned raters who were blind to officer group assignment and time point; a random 20% of the tapes were coded by all three raters for quality assurance. Tapes were coded for global characteristics of the interview (e.g., evocation, collaboration, autonomy-support, empathy), as well as specific behaviors of the officer (e.g., giving information, asking questions, reflecting, confronting). At each time point, we used the mean scores of the two role played interactions. To be consistent with previous studies, we examined two key measures: empathy (a global measure of the extent to which the clinician makes an effort to understand the client's perspective and feelings, scored on a 1–5 scale) and percent MI adherent (an indicator of specific officer behaviors that are consistent with MI, such

as affirming the client, emphasizing personal choice, and asking for permission before providing advice). Interrater reliability (ICCs) for MITI scores was fair for empathy (.539) and good for percent MI adherent (.711) (Cicchetti, 1994; Shrout & Fleiss, 1979).

Measures (Probationer)

To measure probationer outcome, new cases assigned to each officer during a four-month window were followed for six months from the date of their first officer visit. Because we were examining changes in probationer outcome as a function of officer contact, we excluded probationers if they were reassigned or otherwise failed to meet with their assigned officer at least three times during the study period. Excluded probationers represented less than 5% of eligible cases. Because this was an effectiveness evaluation, we relied on two outcome measures routinely gathered by the agency that we thought might be appropriate measures of good versus poor probation progress. Our two main outcome measures included having one or more drug-positive urinalyses or a "poor" probation outcome, as judged by having probation revoked or experiencing a probation violation such as failure to report, a new criminal charge, or evidence of drug use at any time within six months of their first probation officer visit. Probationer outcomes were obtained directly from county and state criminal justice databases, and provided to researchers without personal identifiers.

Statistical Analysis

First, we were interested in whether the officer self-report measures (DRI-R and ORQ) would correlate with MITI assessments of taped interactions. We used repeated measures ANOVA to compare changes in MI skill between training groups adjusting for baseline scores. Second, we used generalized mixed linear modeling (PROC GLIMMIX) to test the effect of training group on probationer outcome adjusting for significant correlates of probationer outcome and clustering of probation officer groups. Third, we used generalized mixed linear modeling tested the effect of MI skill on probationer outcome adjusting for MI skill scores, significant correlates of probationer outcome, and clustering of probation officer groups. We simultaneously adjusted for MI skill scores (empathy, % MI adherent, and ORQ scores) in modeling since all three were correlated with each other. Analyses were performed with SAS 9.2.

RESULTS

The first study aim involved changes in officer skill as a result of the training. During the project period, 6 officers were transferred to different caseloads or

left the department and 1 withdrew from the study due to time constraints, resulting in a final officer count of 23 for purposes of analyses (7 MI+, 7 MI−, 9 SAU). MITI and ORQ scores tended to be highly correlated with each other, but not with DRI-R scores. For instance, ORQ was correlated with empathy ($r = 0.52$, $p < .0001$) and percent MI adherent ($r = 0.50$, $p < .0001$), but not correlated with DRI-R scores ($r = -0.17$, 0.1706). Table 1 shows mean MI scores and one way ANOVA comparisons at each time point for the three training groups. We found differences in empathy, percent MI adherent, and ORQ scores between MI+ and MI−, but not MI− and SAU. DRI-R scores were not significantly different between any groups. Results from repeated measures ANOVA for MI+ versus MI− are shown in Table 2. Significant Group × Time interactions suggested that training improved the MI+ group's empathy and ORQ scores relative to MI− ($p < .05$). We performed the same repeated measures ANOVA tests for difference between MI− and SAU and results were not significant (results not shown) which suggested that interest in MI training alone did not have any effect on MITI scores.

The second study aim involved probationer outcome as a function of the study group. The final sample consisted of 380 probationers. During the four-month entry period, the MI+, MI−, and SAU groups (7, 7, and 9 officers, respectively) were assigned 86, 109, and 185 probationers, respectively. Significant predictors of poor outcomes were used as covariates in

TABLE 1 MI Scores of Probation Officers by Treatment Group

	Supervision as usual ($n = 9$) *M* (*SD*)	MI− (interested) ($n = 7$) *M* (*SD*)	$p^{a,b}$	MI+ (trained) ($n = 7$) *M* (*SD*)	$p^{a,c}$
Empathy					
Baseline	2.28 (0.71)	2.36 (0.69)	0.8259	2.50 (1.00)	0.7611
Post (2 months)	1.78 (0.79)	1.79 (0.76)	0.9841	3.21 (0.86)	0.0063
6 months	1.83 (0.50)	1.79 (0.57)	0.8610	3.50 (0.64)	0.0002
% MI adherent					
Baseline	16.72 (12.43)	25.96 (15.38)	0.2040	37.96 (21.24)	0.2491
Post (2 months)	17.04 (12.29)	23.31 (10.37)	0.3050	62.08 (17.17)	0.0003
6 months	21.35 (10.29)	22.93 (13.56)	0.7952	64.86 (20.20)	0.0007
ORQ					
Baseline	2.23 (1.33)	1.37 (0.29)	0.1159	3.03 (0.80)	0.0003
Post (2 months)	1.87 (0.72)	1.73 (1.11)	0.7683	3.51 (0.61)	0.0030
6 months	1.97 (1.02)	1.50 (1.16)	0.4058	3.77 (0.59)	0.0006
DRI-R					
Baseline	166.33 (17.03)	167.86 (16.41)	0.8595	161.71 (15.97)	0.4913
Post (2 months)	164.89 (21.05)	169.71 (12.92)	0.6035	164.57 (15.81)	0.5177
6 months	167.33 (19.18)	172.29 (18.86)	0.6139	169.29 (12.59)	0.7324

Note. MI = motivational interviewing; ORQ = Officer Responses Questionnaire; DRI-R = Dual-Role Relationships Inventory-Revised.

[a]One-way ANOVA comparisons of means.

[b]Supervision-as-usual versus MI−.

[c]MI− versus MI+.

TABLE 2 Repeated Measures ANOVA Tests for Effects of Group, Time, and Training Group × Time Comparing MI+ Versus MI−

	Training group (between groups tests)		Time (within groups tests)		Training group × Time (within groups tests)	
	Test	*p*-value	Test	*p*-value	Test	*p*-value
Empathy	$F = 12.05$	.0046	$F_2 = 0.47$	.6334	$F_2 = 6.85$	.0044
MI adherent	$F = 36.85$	< .0001	$F_2 = 2.12$	.1420	$F_2 = 3.34$	.0525
ORQ	$F = 33.34$	< .0001	$\Lambda_{2,11} = 0.95$	.4152	$\Lambda_{2,22} = 4.32$	.0413
DRI-R	$F = 0.39$	.5430	$F_2 = 2.13$	.1406	$F_2 = 0.15$	.8612

Note. MI = motivational interviewing; ORQ = Officer Responses Questionnaire; DRI-R = Dual-Role Relationships Inventory-Revised. $\Lambda_{NDF/DDF}$ = Wilks' Lambda test, used when assumption of sphericity not rejected. F_{DF} = F test, used when assumption of sphericity rejected.

subsequent linear modeling: ethnicity, age, supervision level, and charge (results not shown). The odds of having a poor probation outcome were higher for probationers at the medium- or high-risk level ($OR = 2.01$), probationers with felony charges ($OR = 1.63$), and Black probationers ($OR = 1.55$). Specifically: 65% of medium/high supervision level had poor outcomes compared to 49% of low supervision level probationers; 66% of felony had poor outcomes compared to 54% of misdemeanor probationers; 69% of Black probationers had poor outcomes compared to 59% of White probationers. Table 3 shows probationer characteristics by study group. In terms of group equivalency, the MI+ group was assigned significantly higher proportions of probationers that were Black, medium or high risk, and felons.

Table 3 also shows the percent of probationers with a poor outcome or drug-positive urinalysis at six months. The effect of training group on

TABLE 3 Probationer Characteristic/Outcomes by Training Group

	MI+ ($n = 86$)	MI− ($n = 109$)	SAU ($n = 185$)	*p*-value[a]
Male	76.74	66.97	77.30	0.1229
Age (years) *M* (*SD*)	32.49 (11.90)	32.76 (11.33)	31.42 (10.91)	0.5655
Ethnicity				
White	17.44	34.86	36.76	<0.0001
Black	68.60	39.45	32.97	
Hispanic	13.95	24.77	29.19	
Other	0	0.92	1.08	
Supervision level				
Low	12.94	22.02	30.81	<0.0001
Medium	50.59	45.87	56.22	
High	36.47	32.11	12.97	
Felony charge	94.19	58.72	37.30	<0.0001
Poor outcome	67.44	55.05	61.08	0.2110
Drug-positive urinalysis (*SD*)	32.31 (37.41)	33.07 (41.27)	35.90 (41.70)	0.8210

Note. Percentages unless noted. MI = motivational interviewing; SAU = supervision as usual.
[a]Chi-square *p*-value for categorical variables, one-way ANOVA for continuous variables.

TABLE 4 Generalized Linear Modeling of Group on Six-Month Probationer Outcome

	Poor outcome		Drug-positive urinalysis	
	Model 1	Model 2	Model 1	Model 2
MI+	1.692 (0.938, 3.052)	1.211 (0.636, 2.305)	1.014 (0.412, 2.494)	0.898 (0.339, 2.374)
SAU	1.282 (0.792, 2.073)	1.370 (0.816, 2.302)	1.360 (0.613, 3.018)	1.190 (0.504, 2.811)

Note. MI = motivational interviewing; SAU = supervision as usual; Model 1 = MI group (MI− as reference group); Model 2 = MI group + ethnicity, age, supervision level, charge. Odds ratio (95% confidence interval).

probationer outcome was examined via generalized linear mixed modeling (using SAS PROC GLIMMIX). Table 4 shows the effect of study group on our two outcome variables. After controlling for the previously mentioned covariates, group assignment was not associated with the probability of having either a poor outcome or drug-positive urinalysis at six months.

The final study aim was to determine whether probationer outcome was predicted by officer MI skill, independent of study group. Table 5 shows the results of generalized linear mixed modeling of average two- and six-month empathy, percent MI adherent, and ORQ officer scores on 6-month probationer outcome. (DRI-R scores, a self-report measure, were omitted from analyses for having poor relation to other variables and no relation to outcome.) Before adjusting for probationer covariates, we found that officer empathy was related to a more positive probationer outcome. Specifically, a one-point increase in empathy was related to a 29% decrease in odds of a probationer having poor outcome when adjusting for percent MI adherent and ORQ scores. However, after adjusting for probationer covariates, the association was slightly weaker and the confidence interval (CI) was wider and therefore we did not find a significant effect of MI skill on the probability of having a poor outcome or drug-positive urinalysis at six months.

TABLE 5 Generalized Linear Modeling of Officer MI Skill on Six-Month Probationer Outcome

	Poor outcome		Drug-positive urinalysis	
	Model 1	Model 2	Model 1	Model 2
Empathy	0.712 (0.508, 0.999)[a]	0.721 (0.499, 1.042)	0.880 (0.520, 1.487)	0.801 (0.452, 1.420)
% MI Adherent	1.006 (0.990, 1.023)	1.002 (0.985, 1.020)	1.010 (0.988, 1.033)	1.011 (0.988, 1.036)
ORQ	1.252 (0.956, 1.639)	1.206 (0.905, 1.607)	0.740 (0.513, 1.066)	0.775 (0.518, 1.160)

Note. MI = motivational interviewing; ORQ = Officer Responses Questionnaire; Model 1 = Empathy, % MI adherent, ORQ; Model 2 = Empathy, % MI adherent, ORQ, ethnicity, age, supervision level, charge. Odds ratio (95% confidence interval).

[a]Solution for fixed effects $p = .0495$.

DISCUSSION

This study was a randomized trial of MI as a probation supervision strategy. The study was designed to evaluate the effect of a modest MI training program on probation officer skill, on client outcome, and the overall relationship between officer skill and client outcome. The study was strengthened by a randomized design and a relatively rigorous training and evaluation system in a "real world" setting. Major limitations included the lack of group equivalence in terms of ethnicity, charge, and supervision level, as well as the number of new clients assigned to officers during the study period. In general, our MI+ group was assigned a smaller number of more difficult clients. Because of the selection bias inherent in our provider recruitment methods, we made an attempt to balance the research design through randomizing interested providers to trained and untrained group. Unfortunately, did not collect additional information to compare those who expressed interest in receiving training (MI+ and MI−) and those who did not express interest (SAU). This limits our ability to determine how representative the MI+ and MI− groups were of the general probation officer population. However, given the similarity of the groups on baseline MITI scores, it does appear that the three groups were at least similar on initial MI skill. Our power to detect significant differences may have also been hampered by the relatively small number of providers in the study. Indeed, given some of our marginal effect sizes, it is possible that future studies with a larger sample might find a significant effect on offender behavior. In implementing the study, we faced a number of system constraints that are common to large probation departments, including relatively high client loads, difficulty securing resources, and a high officer turnover rate. Finally, we were not able to measure actual changes in officer behavior, instead using standardized role play interactions as a proxy measure of MI competence. We relied on role play interactions because of difficulties in obtaining equivalent practice tapes from the various groups, and because of probationer confidentiality issues. Additionally, because MI is a technique for promoting behavior change, we wanted to assure that the coded samples reflected clients who were struggling with some area of change directly related to their probation progress.

With regard to our first question, we found that an MI training curriculum did improve some key measures of officer MI skill and that these improvements were maintained over six months. Results would also argue against any contamination between officers serving at the same field office; whereas MI+ officers increased their skills, the other two groups showed an overall decrease over the study period. Our training program was designed along the lines of previous research, which has indicated that ongoing supervision and feedback are necessary to produce long-term

changes in provider behavior (Miller, Yahne, Moyers, Martinez, & Pirritano, 2004; Walters, Matson, Baer, & Ziedonis, 2005). At six months, the MI+ group showed modest improvements on two key MI indicators, demonstrating average empathy scores of 3.5 and percent MI adherent scores of 64.8%. As a comparison, the recommended thresholds for beginning proficiency in MI are 3.5 for empathy and 90% for MI adherent (Moyers et al., 2007). This indicates that, although improved over the other groups, MI+ officers still fell below the desired level on at least one key measure. (Other MITI indicators, not discussed, showed modest gains that were mostly at, or just below, minimum competency levels.) However, at the same time, these scores must be interpreted with caution; separate norms have not been developed for probation interactions, which may differ in important ways from the counseling interactions used to establish MITI norms. Interestingly, our two objective measures—the ORQ and MITI—were highly correlated with each other, but unrelated to DRI-R scores. Although the DRI-R measures self-reported helpfulness, rather than MI skills per se, this may also suggest that officers are not good evaluators of their own level of helpfulness and that more objective measures of provider behaviors (like the ORQ and MITI, or even client ratings) are more representative of actual practice. This is consistent with at least one past study that suggested that probationer DRI-R ratings, but not officer ratings, predicted probationer outcome (Skeem et al., 2007).

With regard to our second question, we found that probationer outcome did not differ between training groups on two key outcome measures (i.e., positive urinalysis and poor probation outcome). There are several possible reasons for this lack of effect. For instance, it is possible that MI was adequately delivered in practice but had no effect on the client behaviors we measured. There is good support in the literature for the effectiveness of MI in a range of behavioral outcomes, though the effect of MI as a probation supervision style has not been previously examined. Our outcome measures included only two gross indices of probation progress, but there may have been other behaviors, such as treatment engagement or readiness of change, that were affected. Alternatively, it is possible that, despite modest MI gains on the role play interactions, MI was not adequately delivered in practice. Under this scenario, the lack of difference might be attributed to the training inadequacies or system factors, rather than MI per se. Because we targeted a general probation population, we also had to rely on two general measures of probation progress, rather than measures that might have been targeted to each probationer's specific needs. Finally, as discussed previously, we were unable to sample actual probation interactions due to confidentiality issues and the heterogeneity of clients in a regular supervision caseload. In any case, it is likely that system factors, such as a lack of time with clients, difficulty securing resources,

and high officer turnover, play a large part in probationer outcome in addition to officer supervision style.

With regard to the third question, we found that counselor MI practice, again judged via standardized role play interactions, was mostly unrelated to client outcome. Officer empathy, which was initially related to probationer outcome, was not significant after adjusting for probationer covariates. This suggests again that either MI skill by itself was insufficient to impact client behavior, or that role play competence was unrepresentative of actual practice. Taken together, our results suggest that it is possible to produce changes in probation officer behavior under relatively controlled conditions, but that these changes may be insufficiently linked to client outcome apart from additional systemic changes.

There is a wealth of evidence suggesting that programs that target factors such as substance abuse, changing peer associates, and antisocial beliefs, can strongly impact client behavior (Andrews et al., 1990). Probation system factors, such as a lack of services, ineffective program referrals, and system barriers are also importantly linked to probation outcome (Duffee & Carlson, 1996; Taxman, 1998; Taxman, Perdoni, & Harrison, 2007). Research from fields outside of criminal justice suggests that the provider behavior, apart from technique, plays a large role in outcome (Miller, 2000), however as noted by Bonta, Rugge, Scott, Bourgon, and Yessine (2008) and others, relationship-oriented treatments may be ineffective in the absence of more structured activities that target offender risk and need. In this study, we made no explicit effort to monitor the integration of MI with other evidence-based strategies, and thus future studies might consider implementing MI along with a set of strategies that target risk and need. Future studies might also consider obtaining samples of actual officer/client interactions to determine whether MI skills observed during role play interactions generalize into everyday practice. Finally, in most studies, MI has been examined as a discrete intervention with a single behavioral target. Another role for MI in criminal justice might be as a discrete "pre-treatment" intervention for a specific behavioral outcome (e.g., prior to entering treatment), rather than as an ongoing supervision strategy. Indeed, as we have mentioned, we found a number of factors such as time constraints, high caseloads, and court monitoring duties, that made MI (or potentially any change-focused strategy) very difficult to implement within the current probation system. This study suggested that probation officers can be taught to use MI in standardized probation interactions, though we did not find that such skills were linked to probation outcomes. Although this study provides only limited evidence for the usefulness of MI in a community corrections setting, given the strength of evidence for MI in many similar areas, it would appear that further research into MI in corrections would be warranted in the effort to develop more effective correctional programs.

REFERENCES

Alexander, M., VanBenschoten, S. W., & Walters, S. T. (2008). Motivational interviewing training in criminal justice: Development of a plan. *Federal Probation*, *72*, 61–66.

Andrews, D. A., & Bonta, J. (2003). *The psychology of criminal conduct*. Cincinnati, OH: Anderson Publishing.

Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P., & Cullen, F. T. (1990). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology*, *28*, 369–404.

Baker, A., Lewin, T., Reichler, H., Clancy, R., Carr, V., Garrett, R., . . . Terry, M. (2002). Evaluation of a motivational interview for substance use within psychiatric in-patient services. *Addiction*, *97*, 1329–1337.

Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (vol. 6, pp. 1–62). New York, NY: Academic Press.

Bennett, G. A., Roberts, H. A., Vaughan, T. E., Gibbins, J. A., & Rouse, L. (2007). Evaluating a method of assessing competence in motivational interviewing: A study using simulated patients in the United Kingdom. *Addictive Behaviors*, *32*, 69–79.

Bonta, J., Rugge, T., Scott, T., Bourgon, G., & Yessine, A. K. (2008). Exploring the black box of community supervision. *Journal of Offender Rehabilitation*, *47*, 248–270.

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normal and standardized assessment instruments in psychology. *Psychological Assessment*, *6*, 284–290.

Clark, M., Walters, S. T., Gingrich, R., & Meltzer, M. (2006). Motivational interviewing for probation officers: Tipping the balance towards change. *Federal Probation*, *70*, 38–44.

Daley, D. C., Salloum, I. M., Zuckoff, A., Kirisci, L., & Thase, M. E. (1998). Increasing treatment adherence among outpatients with depression and cocaine dependence: Results of a pilot study. *The American Journal of Psychiatry*, *155*, 1611–1613.

Duffee, D. E., & Carlson, B. E. (1996). Competing value premises for the provision of drug treatment to probationers. *Crime & Delinquency*, *42*, 574–592.

Ginsburg, J. I. D., Mann, R. E., Rotgers, F., & Weekes, J. R. (2002). Motivational interviewing with criminal justice populations. In W. R. Miller, & S. Rollnick (Eds.), *Motivational Interviewing: Preparing people to change* (pp. 333–346). New York, NY: Guilford Press.

Harper, R., & Hardy, S. (2000). An evaluation of motivational interviewing as a method of intervention with clients in a probation setting. *British Journal of Social Work*, *30*, 393–400.

Heather, N., Rollnick, S., Bell, A., & Richmond, R. (1996). Effects of brief counseling among male heavy drinkers identified on general hospital wards. *Drug and Alcohol Review*, *15*, 29–38.

Hettema, J., Steele, J., & Miller, W. R. (2005). Motivational interviewing. *Annual Review of Clinical Psychology*, *1*, 91–111.

Hohman, M., Doran, N., & Koutsenok, I. (2009). Motivational interviewing training for juvenile corrections staff in California: One year initial outcomes. *Journal of Offender Rehabilitation*, *48*, 635–648.

Lundahl, B. W., Tollefson, D., Kunz, C., Brownell, C., & Burke, B. (2010). Meta-analysis of Motivational interviewing: Twenty five years of research. *Research on Social Work Practice*, *20*, 137–160.

McMurran, M. (2002). *Motivating offenders to change: A guide to enhancing engagement in therapy*. West Sussex, England: John Wiley & Sons.

McMurran, M. (2009). Motivational interviewing with offenders: A systematic review. *Legal & Criminological Psychology*, *14*, 83–100.

Miller, W. R. (2000). Rediscovering fire: Small interventions, large effects. *Psychology of Addictive Behaviors*, *14*, 6–18.

Miller, W. R., Hedrick, K. E., & Orlofsky, D. R. (1991). The Helpful Responses Questionnaire: A procedure for measuring therapeutic empathy. *Journal of Clinical Psychology*, *47*, 444–448.

Miller, W. R., Meyers, R. J., & Tonigan, J. S. (1999). Engaging the unmotivated in treatment for alcohol problems: A comparison of three strategies for intervention through family members. *Journal of Consulting and Clinical Psychology*, *67*, 688–697.

Miller, W. R., & Mount, K. A. (2001). A small study of training in motivational interviewing: Does one workshop change clinician and client behavior? *Behavioural and Cognitive Psychotherapy*, *29*, 457–471.

Miller, W. R., & Rollnick, S. (2002). *Motivational interviewing: Preparing people for change*. New York, NY: Guilford Press.

Miller, W. R., Yahne, C. E., Moyers, T. B., Martinez, J., & Pirritano, M. (2004). A randomized trial of methods to help clinicians learn motivational interviewing. *Journal of Consulting and Clinical Psychology*, *72*, 1050–1062.

Moyers, T. B., Martin, T., Manuel, J. K., Hendrickson, S. M., & Miller, W. R. (2005). Assessing competence in the use of motivational interviewing. *Journal of Substance Abuse Treatment*, *28*, 19–26.

Moyers, T. B., Martin, T., Manuel, J. K., & Miller, W. R. (2003). The Motivational Interviewing Treatment Integrity (MITI) code. Unpublished manuscript, University of New Mexico.

Moyers, T. B., Martin, T., Manuel, J. K., Miller, W. R., & Ernst, D. (2007). *Revised global scales: Motivational interviewing treatment integrity 3.0 (MITI 3.0)*.

Project MATCH Research Group. (1998). Matching alcoholism treatments to client heterogeneity: Project MATCH three-year drinking outcomes. *Alcoholism, Clinical and Experimental Research*, *22*, 1300–1311.

Rogers, C. R. (1961). *On becoming a person: A therapist's view of psychotherapy*. Boston, MA: Houghton Mifflin.

Rubak, S., Sandboek, A., Lauritzen, T., & Christensen, B. (2005). Motivational interviewing: A systematic review and meta-analysis. *British Journal of General Practice*, *55*, 305–312.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*, 420–428.

Skeem, J. L., Louden, J. E., Polaschek, D., & Camp, J. (2007). Assessing relationship quality in mandated community treatment: blending care with control. *Psychological Assessment*, *19*, 397–410.

Taxman, F. S. (1998). *Reducing recidivism through a seamless system of care: Components of effective treatment, supervision, and transition services in the community*. Washington, DC: Office of National Drug Control Policy.

Taxman, F. S., Perdoni, M. L., & Harrison, L. D. (2007). Drug treatment services for adult offenders: The state of the state. *Journal of Substance Abuse Treatment*, *32*, 239–254.

Vivian-Byrne, S. (2004). Changing people's minds. *Journal of Social Aggression*, *10*, 181–192.

Walters, S. T., Alexander, M., & Vader, A. M. (2008). The Officer Responses Questionnaire: A procedure for measuring reflective listening in probation and parole settings. *Federal Probation*, *72*, 67–70.

Walters, S. T., Clark, M. D., Gingerich, R., & Meltzer, M. (2007). *Motivating offenders to change: A guide for probation and parole*. Washington, DC: National Institute of Corrections, U.S. Department of Justice.

Walters, S. T., Cornett, M., & Vader, A. M. (in press). Coding motivational interviewing in criminal justice interactions: Recommendations for research and practice. *Motivational Interviewing: Training, Research Implementation, & Practice*.

Walters, S. T., Matson, S. A., Baer, J. S., & Ziedonis, D. M. (2005). Effectiveness of workshop training for psychosocial addiction treatments: A systematic review. *Journal of Substance Abuse Treatment*, *29*, 283–293.

Walters, S. T., Vader, A. M., Harris, T. R., Field, C. A., & Jouriles, E. N. (2009). Dismantling motivational interviewing and feedback for college drinkers: A randomized clinical trial. *Journal of Consulting and Clinical Psychology*, *77*, 64–73.

Woodall, W. G., Delaney, H. D., Kunitz, S. J., Westerberg, V. S., & Zhao, H. (2007). A randomized trial of a DWI intervention program for first offenders: Intervention outcomes and interactions with antisocial personality disorder among a primarily American-Indian sample. *Alcoholism: Clinical and Experimental Research*, *31*, 974–987.

Wormith, J. S., Althouse, R., Simpson, M., Reitzel, L. R., Fagan, T. J., & Morgan, R. D. (2007). The rehabilitation and reintegration of offenders—The current landscape and some future directions for correctional psychology. *Criminal Justice and Behavior*, *34*, 879–892.