Published in final edited form as:

Train Educ Prof Psychol. 2015 May ; 9(2): 175–182. doi:10.1037/tep0000073.

Studying Process and Proximal Outcomes of Supervision for Motivational Interviewing

L.A.R. Stein^{a,b,c,d}, Mary Clair^{a,c}, Shayna Soenksen^a, Rosemarie A. Martin^b, and Jennifer G. Clarke^e

^aUniversity of Rhode Island ^bCenter for Alcohol and Addiction Studies, Brown University ^cRhode Island Training School ^dCenter for Prisoner Health & Human Rights, Brown University ^eMemorial Hospital of Rhode Island, Brown University

Introduction

Supervision process and outcomes are an essential aspect of accountability and effectiveness in clinical supervision but have been a neglected topic. This is a significant omission given potential impact of supervision on client outcomes. This report presents procedures for assessing training outcomes and quality of ongoing supervision process as rated by supervisee and supervisor. Although these procedures can be utilized across treatment approaches, this paper focuses on supervision with respect to Motivational Interviewing (MI; Miller & Rollnick, 2013), a treatment widely disseminated for behavior change (Martino, Carroll, & Ball, 2007). Central to use of MI are therapist empathy, eliciting change talk, focusing on discrepancy between client behaviors and values, encouraging confidence to change, and non-confrontational response to resistance (Moyers, Martin, Manuel, Hendrickson, & Miller, 2005).

Competency-based supervision

Supervision involves observation; encouraging self-assessment; instruction, modeling and mutual problem-solving; feedback and evaluation; and role-modeling (Kaslow, Falender, & Grus, 2012). Supervision has multiple functions including clarifying goals and tasks; promoting a context that supports learning; fostering competence and self-efficacy via recognition of strengths; promoting accountability and science-informed practice; and monitoring and offering feedback regarding quality of service delivery (Kaslow et al., 2012). Competency-based supervision is geared towards identifying components that make up particular clinical competencies, and strategies to develop such competencies (Falender & Shafranske, 2007). Core competencies and the specific knowledge, skills and values that form them, create a standard that may be applied to supervisors (Falender & Shafranske, 2007). Within the context of competency-based supervision, it is important to develop methods to assess elements of the training and supervision process. Questionnaires offer an empirical way to assess coverage and quality of training and supervision, and systematize

the process across various supervisors, which is important for quality of care. To date, this area has been sorely lacking, relative to tracking both client outcomes and counselor skill development via fidelity instruments.

Various models exist to implement best practices of competency-based supervision, including psychotherapy-based approaches (Falender & Shafranske, 2007). Psychotherapy-based supervision provides a coherent approach to therapy in which knowledge, theory, and technique derived from a specific orientation inform the conduct of treatment and provide a clear focus for supervision; techniques used in therapy are tailored and used as learning strategies in supervision, providing the trainee with direct experience in use of an intervention (Beck, Sarnat, & Barenstein, 2008).

Psychotherapy-based approach applied to MI

Motivational Interviewing (MI) is widely disseminated (Miller & Rose, 2009), and the demand for professional training in it has grown steadily (see Soderlund, Madson, Rubak, & Nilsen, 2011). It is used by an assortment of service providers in a variety of settings (Lundahl & Burke, 2009; Rubak, Sandbaek, Lauritzen, & Christensen, 2005) for a varied number of problem-behaviors from substance use to diet (Miller & Rose, 2009), and is taught in a number of settings including social work, medical, and psychology internship and post-doctoral programs (Baker, Mendoza-Newman & Cornish, 2014; Madson, Loignon, & Lane, 2009). The widespread use of this evidence-based approach speaks to its importance for persons learning the craft of providing therapy, and to professionals tasked with teaching and supervising MI.

To date, there is little in the empirical literature regarding the style to be used by expert trainers or supervisors during dissemination of MI to trainees. In their review of MI training processes, Madson et al. (2009) articulate components of training, but no recommendations are provided about trainer or supervisor use of MI during interactions with trainees. It is relatively rare to find studies articulating how supervisors are trained or what supervisors are to do during supervision (Baer et al., 2007). Indeed the importance of conducting MI training and supervision in a manner consistent with MI (i.e., using a psychotherapy-based approach) has been underscored (Martino et al., 2007; Kenyon, 2007), but to date there have been no empirical studies on the topic. Therefore, a psychotherapy-based approach was applied in which knowledge, theory, and technique derived from MI informed training, supervision and the conduct of treatment. Further, techniques used in MI were tailored and used as learning strategies in training and supervision (e.g., eliciting from trainees their own reasons for improved practice, which is similar to eliciting from clients their own reasons for change).

The current study

In an effort to evaluate the use of this approach in learning MI, this study undertook several aims: **Aim 1**) Measure pre- and post-training MI-skills (proximal outcomes); improvement in skill suggests that the training approach was effective. **Aim 2**) Measure the relationship between post-training skill and conduct of training (coverage of MI knowledge, theory and skills, and use of MI techniques to provide training); positive relationship between post-

training skill and training ratings is consistent with the idea that the training approach was effective. **Aim 3)** Determine the extent to which a psychotherapy-based approach based on MI was applied in both training and ongoing supervision; significant relationship between training and supervision ratings suggests that a uniform approach was applied. **Aim 4)** Determine the extent to which supervisor experience might impact training and supervision; supervisors more experienced in a particular therapy should apply psychotherapy-based supervision based on that therapy more effectively. **Aim 5)** Determine the extent to which this psychotherapy-based approach to training and supervision is applicable across trainee types; little or no differences is consistent with the idea that these procedures are applicable across trainee types.

Methods

Data are from an active lab that trains graduate students and post-doctoral fellows from four universities, and that selects and trains research staff involved in clinical intervention studies. This is an archival study examining use of a psychotherapy-based approach to training, evaluating and supervising trainees in MI. Data were pooled among student trainees and research staff to increase sample size, because methods used for trainees and research staff were similar, and to evaluate the applicability of the approach across learners.

Participants

The sample ($N = 36$) included 10 graduate trainees in an academic setting who wanted to learn behavioral intervention, including MI. It also included 26 job applicants or other research staff who participated as part of employment in conducting funded research in behavior change strategies, including MI. The sample participated from 2007 to 2012. For job applicants not hired ($n=11$), detailed data on age, degree and race were not retained; however, we estimate that over 90% were White, 6% were Hispanic, over 80% were female, over 90% had at least a bachelor's degree, and ages ranged from approximately 23 to 40 years.

The sample of student trainees ($n=10$), research staff ($n=5$) and hired job applicants ($n=10$) is described as follows: 8% Hispanic, 24% Non-White, 84% female, 40% had a bachelor's degree, 48% had a masters degree, and mean (M) age was 30.72 years (standard deviation, $SD = 9.45$ years). The sample used in analyses of study hypotheses consisted of 10 students, 5 research staff, and 10 hired job applicants, all of whom received MI training and supervision. Un-hired job applicants were only included in inter-rater reliability analyses.

Procedures

Training was offered in use of MI to interested students (graduate and post-doc) from four local universities. Students were eligible if they completed at least one year of graduate work and committed to 5-10 hours/week of training for 3-9 months. Training was conducted to facilitate supervisee work with clients found in university and correctional settings. Research staff and job applicants participated as part of employment on four funded grants studying behavior change in adult and juvenile settings including correctional, school, and

social service. Institutional Review Board (IRB) approval was obtained for de-identified archival data analysis.

Trainees (students, research staff and job applicants) were required to take the Helpful Responses Questionnaire (HRQ; Miller, Hedrick, & Orlofsky, 1991) prior to training (see Measures). Because beginning skill-level is related to MI performance over time (Carpenter et al., 2012), HRQ pre-test was used as part of selection procedures for job applicants. Following pre-test, training sessions in MI were provided, after which a training evaluation was administered, along with HRQ post-test. After several months (*Median* = 5 months) of providing MI with coached supervision, an evaluation of supervision was conducted. Training and supervision evaluations took place privately and confidentially, with identifiers removed to encourage honest responding. Participants were told that supervisors would examine results after several were amassed over time in order to foster anonymity. Participants were asked to create a unique identifier using letters of their mother's maiden name and two digits of their social security number.

Training description—Training approach was flexible, and delivered in a manner consistent with MI. A common model of training consists of two days of didactic training, role-plays, and feedback with expert coaching (Martino et al., 2007). When training students, due to conflicting schedules, coursework, and the like, a two-day training model was not always optimal. Therefore, training was broken into several 1-2 hour segments per week spread over several weeks to obtain the equivalent of the above two-day training. When training research staff, depending on the study, traditional two day, segmented, or more intensive training was employed. Intensive training consisted of about five full days of training including didactics, role-plays, feedback and coaching. More time was taken to model MI, engage in role-plays and feedback, and receive expert coaching. In the event that treatment was manualized, portions of the script were utilized during role-plays. No matter format (two-day, segmented, or intensive), training was provided as outlined in seminal sources (Miller & Rollnick, 1991, 2002, & 2013). A primary goal was to promote trainee ability to respond empathically and strategically to client change- and counter-change statements (Miller, Yahne, Moyers, Martinez, & Pirritano, 2004). Readings were provided (Center for Substance Abuse Treatment, 1999; Miller & Rollnick, 2002), videos utilized (Miller, Rollnick, & Moyers, 1998), and hand-outs reviewed (e.g., 12 road blocks to effective listening; see Gordon, 1970). Criteria utilized in determining MI skill level (using the Motivational Interviewing Treatment Integrity Scale (MITI); Moyers, Martin, Manuel, Miller, & Ernst, 2010) was also reviewed. Training usually consisted of about 2-6 persons at a time, and was provided by the below described supervisors.

Supervision description—Two supervisors provided supervision for this study. Both had doctoral degrees from programs accredited by the American Psychological Association. They had fifteen and seven years of MI experience between them, and twelve and five years of experience in MI supervision. Both received training modeled after the MI Network of Trainers (MINT) model, and MITI training from Dr. Theresa Moyers (see Moyers et al., 2010).

Supervision was delivered using MI style. Supervisors began by eliciting trainee point of view with open-ended questions (“What are your thoughts on the session?”), and then reflections to acknowledge effort, successes, and challenges (“You felt well-prepared for the session and were easily able to elicit her reasons for change; on the other hand, she became quiet suddenly when you offered your concerns.”). Supervisors provided MITI feedback, and then solicited trainee response to feedback starting with relative strengths. Trainee successful efforts to use MI were affirmed (“Your use of decisional balance was well-timed; I’m wondering what made you think of using it at that point.”), and supervisors worked collaboratively with trainees to problem-solve and improve MI delivery (“You might have maintained a more collaborative spirit by asking permission before sharing your concerns for her safety.”). By acknowledging and affirming strengths, the intent was to enhance trainee self-efficacy. Supervisors solicited from trainees examples of good and not-so-good MI practice and impact on the client. Through this process, the intent was that trainees learn to recognize and respond skillfully to client change- and counter-change language. Supervisors communicated empathy as well (“It’s not always easy to learn a new skill;” “You’re pleased with how well the client received your summary.”). Supervisors provided a selective summary of the trainee’s supervision session, including strengths, and asked what the trainee might add or change to the summary. Supervisors then asked what trainees would like to do differently, and collaborated on setting practice goals for the next session and what might assist in meeting goals.

Typically, group supervision was utilized. Supervisors began by summarizing the agenda set in collaboration with members during the last meeting, and asked how members might add to or alter the summary. Trainees got about 10-20 minutes of time devoted directly to their practice. Group members were actively engaged in generating observations, questions, suggestions and role-play. They could volunteer a 5-10 minute segment for which they wanted assistance from group members and supervisor, or a segment that went particularly well. They sometimes engaged in MITI ratings to compare to supervisor and other group members. Supervisors expressed appreciation for trainee willingness to be open during group supervision. Supervisors provided a brief summary of the group session, acknowledged goals that were set, and solicited input on agenda for the next group session. Supervision was weekly or biweekly, spanned 60-90 minutes and consisted of 2-6 trainees.

Measures

Helpful Responses Questionnaire (HRQ; Miller et al., 1991)—This is a brief, free-response questionnaire designed to measure reflective listening and empathic ability, skills which are central to the implementation of MI. Six brief scenarios are presented in which a troubled person presents a problem and the participant then writes down 1-2 responses that would be most helpful. Each response is rated on a 5-point scale for depth of reflection and ability to avoid communication roadblocks (1=No reflection plus a comment that can interrupt the flow of conversation [roadblock], 5=A high quality reflection with no roadblock). Miller et al. (1991) found item-level inter-rater reliability coefficients in the .71-.91 range, total score reliability of .93, that scores of community paraprofessionals rose significantly following a 2-day workshop on active listening and crisis intervention skills, and that scores were associated with increased client motivation for change. A low score

(rating of 1 or 2) on an item can be used to diagnose roadblocks to being helpful (for example, changing the subject or giving orders). Following training, both HRQ scores and MI skills (rated during interactions with standardized patients) have been found to significantly improve (Baer et al., 2004).

Training Evaluation—This measure was informed by Baer et al. (2007) and Miller, Sorenson, Selzer, & Brigham (2006) who observed that treatment training often consists of didactic portions for background, review of specific procedures, and demonstrations via video or role-play. Preparatory knowledge is acquired via reading, verbal instruction or observation; followed by practice with expert feedback. Training evaluation consists of 33 items assessing these dimensions, along with trainee impressions of trainer expertise and general training quality. MI-specific items are based on Miller & Rollnick (1991, 2002, & 2013). Items are rated using 4-point Likert scale (1=strongly disagree to 4=strongly agree; or 1=not covered to 4=very useful). Items were sorted by two raters experienced in MI supervision into the following categories: MI-Adherent (3 items; for example, Trainer was respectful), Trainer quality (3; Trainer has expertise), Training quality (4; Training was effective), Information (3; Cultural issues were raised), Practice skills (3; I had ample opportunity to practice skills), Modeling (3; Trainer demonstrated MI in role-plays), and Directiveness (14; Types of reflections were covered). Agreement rate was 75%; and disagreements regarding sorting of items were resolved via discussion.

Supervision Evaluation—Ongoing coaching appears to be important in learning an intervention (Miller et al., 2006), and in particular sustained coaching has been found to be critical in learning MI (Miller et al., 2004). Presumably, coaching should reinforce philosophy and skills presented during initial training. In addition, consistent with psychotherapy-based supervision, MI may be utilized during aspects of supervision delivery to further model use and to reduce trainee ambivalence about learning a new intervention (Kenyon, 2007; Martino et al., 2007). Supervision was evaluated using 35 items rated on a 6-point Likert scale (1=not at all to 6=extensively). Items were selected based on Martino et al. (2007) and Kenyon (2007). Items were sorted by two raters experienced in MI supervision into the following categories: MI-Adherent (13 items; for example, Supervisor affirmed counselor successful efforts), Supervision quality (1; Supervision was what a counselor should expect), Coaching (4; Supervisor provided feedback on MI implementation using a rating system), Evocation (6; Supervisor asked key questions to renew counselor commitment), Collaboration (5; Supervisor developed specific coaching tips with counselor), Autonomy/support (3; I know why I want to use MI in treatment), and Empathy (3; Supervisor acknowledged counselor challenges). Agreement rate was 69%; and disagreements regarding sorting of items were resolved via discussion.

Analyses

Scores were examined for distributional assumptions. Supervisors double-rated over 20% of pre- and post-test scores to assess reliability (percent agreement used due to small *N*). Agreement was defined as being within +1 per item. For job applicants, pre-test status was frequently known, since results were used as part of screening process. However, for double coding, the second rater was blind to pre- and post-test status. To examine **aim 1** (measure

pre- and post-training MI-skill), HRQ pre- and post-test difference scores were evaluated using dependent-tests; and post-test scores were utilized to identify common errors that remained even after selection and training.

For **aim 2** (measure the relationship between post-training skill versus training coverage of MI knowledge, theory and skills, and use of MI techniques to provide training), training session evaluation scores were correlated with post-test HRQ scores. To evaluate **aim 3** (determine the extent to which a psychotherapy-based approach based on MI was applied in both training and ongoing supervision), correlations were performed to determine if those rating training sessions more favorably perceived more MI-consistent supervision activities.

For **aim 4** (determine the extent to which supervisor experience might impact training and supervision) differences between trainees associated with each supervisor were examined in terms of HRQ post-tests, training session scores, and supervision rating scores using independent t-tests. Finally, to examine **aim 5** (determine the extent to which this psychotherapy-based approach to training and supervision is applicable across trainee types), differences between student trainees and research staff were evaluated in terms of HRQ post-tests, training session scores, and supervision rating scores using independent t-tests.

Results

Standardized training has been used since 2000, based on MINT training. In about 2006, the HRQ was included, followed by training evaluation in 2009 and supervision evaluation in 2011. These efforts were begun on funded research projects and later expanded to training students. As a result, sample size differs depending on analyses.

Data were checked for distributional assumptions and no transformations were needed. Total scores were used on all scales and subscales. On average, raters agreed on 4.9 items out of 6 on the HRQ. Raters agreed on total scores on 75% of double-rated HRQs. In evaluating **aim 1**, HRQ pre- and post test scores were significantly different: $t(16)=2.82, p=.012, d=0.71, M_{pre}=12.94, SD_{pre}=6.36, M_{post}=17.53, SD_{post}=6.51$. Across 31 HRQ pre-tests, in total 159 roadblocks were counted covering all 12 types of roadblocks. The most frequently found roadblocks were “questioning or probing,” 49%; “advising, suggesting, or providing solutions,” 23%; and “persuading with logic, arguing or lecturing,” 5%. Only two types emerged at HRQ post-test. In total, 26 roadblocks were counted across 17 HRQ post-tests with 96% and 4% classified as “questioning or probing” and “advising, suggesting or providing solutions,” respectively.

In evaluating **aim 2**, HRQ-post test scores correlated with training evaluation scores as follows ($n=12$): $r_{Total}=.65 (p=.021)$, $r_{Modeling}=.63 (.028)$, $r_{Information}=.62 (.031)$, $r_{Directiveness}=.62 (.032)$, $r_{Training\ Quality}=.52 (.082)$, $r_{Practice\ Skills}=.36 (.25)$, $r_{Trainer\ Quality}=.74 (.006)$, $r_{MI-Adherent}=.75 (.005)$. For **aim 3**, correlations between training and supervision evaluations are located in Table 1. Most correlations were statistically significant and produced effect sizes in the large range (see Cohen, 1988), with the exception of the coaching subscale of the supervision evaluation, which generally produced lower non-

significant correlations. Differences between supervisors on HRQ post-test and training evaluation scores ([aim 4](#)) are presented in Table 2. Although sample size mitigated power, effect sizes were in the large range, suggesting differences on outcomes by supervisor. Differences between supervisors on the supervision evaluation are not presented to save space; and again, although results were non-significant due to sample size, effect sizes were in the large range. The journal makes these results available via electronic portal, or readers can contact the first author with inquiries. Similarly for [aim 5](#), HRQ post-tests and training evaluations were compared between research staff and student trainees (Table 3), showing research staff had better outcomes than student trainees. Again, differences between trainee type on the supervision evaluation are not presented to save space. Results were non-significant and effect sizes were in the small range. The journal makes these results available via electronic portal, or readers can contact the first author with inquiries.

Discussion

This study offers preliminary evidence for a psychotherapy-based approach to disseminate MI, supported by measures of process and proximal outcomes as rated by trainees and supervisors. Within the context of competency-based supervision (see Falender & Shafranske, 2007; Kaslow et al., 2012), it is important to develop methods to assess elements of the training and supervision process. To date, this area has been sorely lacking, relative to tracking both client outcomes and counselor skill development via fidelity instruments. Similarly, Baer et al. (2007) note the lack of data regarding supervision and training processes, and in particular, Madson et al. (2009) point to the lack of detail surrounding assessment of training and supervision outcomes with respect to MI. These procedures and measures merit further research with respect to MI in particular, but also to advance supervision and training provided by psychologists more generally.

We are aware of no other published studies applying competency-based supervision using a psychotherapy-based approach to learn MI. Having models and methods to disseminate MI is important given MI's prominence in behavior change (Madson et al., 2009). Findings provide preliminary support for the applicability of competency-based supervision using a psychotherapy-based approach in disseminating MI. Using procedures presented here, the HRQ can be used to evaluate training and further diagnose areas in need of attention to enhance MI skill. We recommend training sessions be conducted using an MI-consistent style and modeled after MINT training as described above. Although not tested here, it stands to reason that training evaluation scores could be used to diagnose perceived deficits in training delivery. We also recommend that supervision should model MI delivery and utilize objective feedback of trainee performance (e.g., using the MITI). Again, although not tested here, supervision evaluation scores could be used to diagnose perceived deficits in modeling MI and maintain good supervision practice.

Measures of training inform what is communicated to trainees, how to improve training, and standardize the training process. Measures of supervision allow supervisors to better track effectiveness of supervision from trainee perspective and standardize the supervision process. Such standardization and tracking are important to clinical practice and research programs using multiple trainers, training sessions, and supervisors over time. These

measures of training and supervision have yet to be more rigorously validated. The item pool may need to be refined with additions and deletions. At this time however, the training and supervision evaluations have some degree of face-validity in that item content during training seems to reflect keeping sessions on track (directiveness), whereas item content during supervision appears to emphasize MI-spirit (MI-A), as might developmentally be expected. Although psychotherapy-based approaches to supervision do not necessarily employ measures of process and proximal outcomes, our use of measures is consistent with other areas in the supervision literature that emphasize measurement (Milne, 2008). We attended to two fundamental and complementary approaches to the evaluation of supervision: Supervisor performance as rated by trainee, and trainee performance as rated by supervisor (see Milne, 2008). We recommend using quantitative evaluation to track coverage and quality of training and supervision.

This paper is relevant to psychologists concerned with education and training for several reasons. First, the sample consists of 8% PhD-, 48% MA-, 40% BA-, and 4% HS-level persons; 60% had some previous exposure to behavioral intervention techniques; 8% were Hispanic, 24% were non-White; and 84% were women. Therefore, in terms of education and experience, the sample of trainees is not unlike those that might be encountered in doctoral, internship and post-doctoral programs. The sample is also similar to those found in training programs in terms of gender, race and ethnicity (American Psychological Association, n.d.). Second, supervisors may wish to utilize an MI style during supervision, even when trainees are not necessarily employing MI in practice, since MI may assist to maintain trainee motivation, bolster trainee confidence to attempt new skills, facilitate trainee acceptance of feedback, and enhance trainee engagement in skill development (Kenyon, 2007; Sobell, Manor, Sobell, & Dum, 2008). More work is needed to test the use of MI in this manner. Third, while content of the measures presented here was specific to MI, constructs assessed by the scales may be applicable more broadly to other forms of supervision, which recognize the importance of empathy, information, modeling, feedback, increasing autonomy over time, and collaboration in supervision. Fourth, training program activities frequently include enhancement of therapeutic skills. While more work is needed to determine generalizability, results suggest that a psychotherapy-based approach to training and supervision assists with trainee skill development. Finally, it is important for training programs to track outcomes, and this study demonstrates how measurement might be used to evaluate skill development and quality of training and supervision from trainee and supervisor perspectives.

With respect to **aim 1**, on average, about 5.1 road blocks were found per HRQ at pre-test, but after selection and training only about 1.5 were found per HRQ at post-test. The most common types of road blocks found at both pre- and post-HRQ tests suggest that “probing” and “providing solutions” are rather refractory and deserve further attention during supervision. Analyses indicate significant improvement in HRQ scores at post-test following training, and that training evaluation scores were positively correlated with HRQ post-test scores (**aim 2**), as expected, with effect sizes generally in the large range. Of note, the scale most associated with trainer use of MI techniques during training (MI-A) was most highly correlated with development of MI-skills following training (HRQ post-test score), consistent with use of a psychotherapy-based approach. This suggests that the training

approach was effective and that supervisors might consider this approach when working with trainees learning MI. More broadly, this suggests supervisor modeling of skills promotes trainee skill development.

Correlations among training and supervision evaluation scales were generally in the large range and significant ([aim 3](#); Table 1). Training and supervision perceptions were expected to be related because trainer and supervisor were often the same person, and a standardized psychotherapy-based approach modeled after MI was used in both training and supervision across trainers and over time. However, the coaching subscale of the supervision evaluation correlated poorly with training subscales. This is noteworthy as a large part of supervision is coaching and providing feedback. It may be that the item pool for the coaching scale needs to be improved.

HRQ post-test, and training and supervision evaluation scores were compared between supervisors ([aim 4](#); Table 2). Although few results were significant due to being underpowered, differences were in the expected direction and effect sizes were generally large, implying that more experienced supervisors may be more effective in assisting trainees learn skills. Ideally no differences would have been found when comparing HRQ post-test, and training and supervision evaluations between research staff and students ([aim 5](#); Table 3). However, unexpected differences were found with research staff having higher scores on the HRQ post-test, and on training evaluation total, directiveness and training quality scales. It is unclear why this occurred, but this may be related to both uneven distribution of students and research staff among supervisors and small sample size. It may also be because research staff members were selected for their pre-test HRQ performance, and because students frequently had modified training format owing to course schedule (see Methods). On the other hand, differences between research staff and students on supervision evaluation scores were insignificant, produced small effect sizes, and direction of differences was inconsistent, as expected. This suggests that the training and supervision approach may have broader applicability across trainee types, but more rigorous work is needed in this area.

There are several limitations to this study. The sample size was small, and this is particularly true for some analyses presented. Replication is obviously in order with larger sample sizes. Effect sizes are offered in an effort to offset small sample size. Because this is a relatively understudied area and because it is relatively difficult to obtain large samples of trainees, these preliminary data are offered in an effort to advance the field and generate further research in the area. Similarly, only two supervisors could be studied, and future research would benefit by including larger numbers of supervisors. It should be noted, however, that it is not unusual to see published works including analyses of only two supervisors (e.g., Sobell et al., 2008). Larger samples would allow for alpha corrections during analyses. Furthermore, larger samples would permit needed refinement of the item pool via principle components analyses and confirmatory factor analyses.

In addition, the data were collected as part of an emerging practice with trainees, as compared to being collected as a part of a well-controlled trial. As a result there was wide heterogeneity with respect to trainees (student vs. research staff), timing of measures was

not standardized, and simple pre-post tests were utilized (as compared to random assignment). On the other hand, creation of procedures and measures in this manner may facilitate dissemination to real world contexts.

Future studies may wish to empirically link training and evaluation scores to client outcomes and trainee fidelity scores on a tool such as the MITI. This study did not address whether use of MI during training and supervision is necessary or whether it is better than other approaches to supervision when learning MI or other interventions. This can be studied in future endeavors as well.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by grants from the National Institute on Drug Abuse, R01 DA-13375 (PI, Stein), R01 DA-018851 (PI, Stein), R01 DA-020731 (PI, Stein), and by grants from the National Institute of Child Health and Human Development, R01 HD-065942 (PI, Stein) and R01 DA-024093 (PI, Clarke).

References

- American Psychological Association. n.d.Retrieved August 14, 2014 from www.apa.org/ed/accreditation/about/research/report-data.aspx
- Baer JS, Ball SA, Campbell BK, Miele GM, Schoener EP, Tracy K. Training & fidelity monitoring of behavioral interventions in multi-site addictions research: A review. *Drug & Alcohol Dependence*. 2007; 87(2-3):107–118. [PubMed: 17023123]
- Baer JS, Rosengren DB, Dunn CW, Wells EA, Ogle RL, Hartzler B. An evaluation of workshop training in motivational interviewing for addiction and mental health clinicians. *Drug & Alcohol Dependence*. 2004; 73(1):99–106. [PubMed: 14687964]
- Baker, J.; Mendoza-Newman, M.; Cornish, J. Internship and Postdoctoral Programs in Professional Psychology. 40th Edition. The Association of Psychology Postdoctoral and Internship Centers; Houston, Texas: 2014. Retrieved August 14, 2014 from www.appic.org
- Beck, J.; Sarnat, J.; Barenstein, V. Psychotherapy-Based Approaches to Supervision. In: Falender, CA.; Shafraanske, EP., editors. *Casebook for clinical supervision: A competency-based approach*. American Psychological Association; Washington, D.C.: 2008. p. 57-96.
- Carpenter KM, Cheng WY, Smith JL, Brooks AC, Amrhein PC, Wain RM, Nunes EV. “Old dogs” and new skills: How clinician characteristics relate to motivational interviewing skills before, during, and after training. *Journal of Consulting and Clinical Psychology*. 2012; 80(4):560–573. [PubMed: 22563640]
- Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed.. Lawrence Erlbaum Associates, Inc.; Hillsdale, NJ: 1988.
- Center for Substance Abuse Treatment. *Enhancing motivation for change in substance abuse treatment*. US Government Printing Office; Washington, D.C.: 1999. Treatment Improvement Protocol Number 35. DHHS Publication No. (SMA) 99-3354
- Falender CA, Shafraanske EP. Competence in competency-based supervision practice: Construct and application. *Professional Psychology: Research and Practice*. 2007; 38(3):232–240.
- Gordon, T. *Parent effectiveness training: The “no-lose” program for raising responsible children*. Peter H. Wyden; New York, NY: 1970.
- Kaslow NJ, Falender CA, Grus CL. Valuing and practicing competency-based supervision: A transformational leadership perspective. *Training and Education in Professional Psychology*. 2012; 6(1):47–54.

- Kenyon, R. Supervising in the style of motivational dialogue. In: Tober, G.; Raistrick, D., editors. *Motivational dialogue: Preparing addiction professionals for motivational interviewing practice*. Routledge; New York, NY: 2007. p. 114-133.
- Lundahl B, Burke B. The effectiveness and applicability of motivational interviewing: A practice-friendly review of four meta-analyses. *Journal of Clinical Psychology: In Session*. 2009; 65(11): 1232-1245. [PubMed: 19739205]
- Madson MB, Loignon AC, Lane C. Training in motivational interviewing: A systematic review. *Journal of Substance Abuse Treatment*. 2009; 36:101-109. [PubMed: 18657936]
- Martino, S.; Carroll, KM.; Ball, SA. Teaching, monitoring and evaluating motivational interviewing practice. In: Tober, G.; Raistrick, D., editors. *Motivational dialogue: Preparing addiction professionals for motivational interviewing practice*. Routledge; New York, NY: 2007. p. 87-113.
- Miller WR, Hedrick KE, Orlofsky DR. The Helpful Responses Questionnaire: A procedure for measuring therapeutic empathy. *Journal of Clinical Psychology*. 1991; 47(3):444-8. [PubMed: 2066417]
- Miller, WR.; Rollnick, S. *Motivational interviewing: preparing people to change addictive behavior*. Guilford Press; New York: 1991.
- Miller, WR.; Rollnick, S. *Motivational interviewing: preparing people for change*. 2nd ed.. Guilford Press; New York: 2002.
- Miller, WR.; Rollnick, S. *Motivational interviewing: Helping people change*. 3rd ed.. Guilford Press; New York: 2013.
- Miller, WR.; Rollnick, S.; Moyers, TB. *Motivational interviewing: Professional training VHS Videotape/DVD Series*. Center on Alcoholism Substance Abuse, and Addictions; The University of New Mexico: 1998. Director
- Miller WR, Rose G. Toward a theory of motivational interviewing. *American Psychologist*. 2009; 64(6):527-537. [PubMed: 19739882]
- Miller WR, Sorenson JL, Selzer JA, Brigham GS. Disseminating evidence-based practices in substance abuse treatment: A review with suggestions. *Journal of Substance Abuse Treatment*. 2006; 31:25-39. [PubMed: 16814008]
- Miller WR, Yahne CE, Moyers TB, Martinez J, Pirritano MJ. A randomized trial of methods to help clinicians learn motivational interviewing. *Journal of Consulting and Clinical Psychology*. 2004; 72(6):1050-1062. [PubMed: 15612851]
- Milne, D. Evaluating and enhancing supervision: An experiential approach. In: Falender, CA.; Shafranske, EP., editors. *Casebook for clinical supervision: A competency-based approach*. American Psychological Association; Washington, D.C.: 2008. p. 211-233.
- Moyers TB, Martin T, Manuel JK, Hendrickson SML, Miller WR. Assessing competence in the use of motivational interviewing. *Journal of Substance Abuse Treatment*. 2005; 28:19-26. [PubMed: 15723728]
- Moyers, TB.; Martin, T.; Manuel, JK.; Miller, WR.; Ernst, D. *Revised Global Scales: Motivational Interviewing Treatment Integrity 3.1.1*. Center on Alcoholism, Substance Abuse and Addictions; University of New Mexico: 2010. Unpublished manual Retrieved from http://www.motivationalinterview.org/Documents/miti3_1.pdf
- Rubak S, Sandbaek A, Lauritzen T, Christensen B. Motivational interviewing: A systematic review and meta-analysis. *British Journal of General Practice*. 2005; 55:305-312. [PubMed: 15826439]
- Sobell L, Manor H, Sobell M, Dum M. Self-critiques of audiotaped therapy sessions: A motivational procedure for facilitating feedback during supervision. *Training and Education in Professional Psychology*. 2008; 2(3):151-155.
- Soderlund L, Madson M, Rubak S, Nilsen P. A systematic review of motivational interviewing training for general health care practitioners. *Patient Education and Counseling*. 2011; 84:16-26. [PubMed: 20667432]

Table 1
Correlations between Training and Supervision Evaluations (n=12)

Training		Supervision							
		Tot	Emp	Coach	Aut	SupQ	Coll	Evoc	MI-A
Tot	<i>r</i>	.84	.71	.24	.73	.77	.68	.65	.88
	<i>p</i>	.001	.009	.451	.007	.004	.016	.024	.001
Model	<i>r</i>	.75	.73	.20	.71	.53	.74	.55	.71
	<i>p</i>	.005	.007	.530	.009	.076	.006	.067	.009
Info	<i>r</i>	.81	.70	.13	.73	.81	.53	.72	.87
	<i>p</i>	.002	.012	.690	.007	.001	.006	.008	.001
Dir	<i>r</i>	.76	.60	.21	.69	.68	.62	.56	.83
	<i>p</i>	.004	.040	.510	.013	.015	.033	.058	.002
TrainQ	<i>r</i>	.83	.61	.30	.62	.85	.61	.66	.88
	<i>p</i>	.001	.017	.338	.032	.001	.036	.020	.001
PSkills	<i>r</i>	.72	.73	.18	.55	.89	.44	.47	.83
	<i>p</i>	.008	.007	.580	.062	.001	.152	.122	.001
TrainerQ	<i>r</i>	.86	.67	.50	.65	.62	.76	.63	.85
	<i>p</i>	.001	.017	.097	.023	.030	.004	.030	.001
MI-A	<i>r</i>	.61	.58	.01	.63	.37	.78	.50	.55
	<i>p</i>	.035	.046	.969	.029	.232	.003	.099	.066

Note: Tot = Total, Emp = Empathy, Coach = Coaching, Aut = Autonomy, SupQ = Supervision quality, Coll = Collaboration, Evoc = Evocation, MI-A = Motivational Interviewing adherent, Model = Modeling, Info = Information, Dir = Directiveness, TrainQ = Training quality, PSkills = Practice skills, TrainerQ = Trainer quality.

Table 2
Supervisor Differences on HRQ and Trainer Evaluation

	Supervisor-1		Supervisor-2		Statistic			
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>df</i>	<i>p</i>	<i>d</i>
HRQ _{po} ^a	20.00	4.84	15.33	7.26	1.54	15	.145	0.76
T-Eval ^b								
Tot	125.88	5.77	111.33	13.56	2.47	6.37 ^c	.047	1.40
Model	12.00	0.01	10.67	1.86	1.75	5.00 ^c	.140	1.01
Info	11.00	1.07	9.17	2.32	1.80	6.60 ^c	.117	1.01
Dir	52.13	2.59	45.17	5.38	3.22	12.00	.007	1.65
TrainQ	15.50	1.07	13.33	2.07	2.35	7.01 ^c	.051	1.32
PSkills	11.50	1.07	10.67	1.21	1.37	12.00	.197	0.73
TrainerQ	11.88	0.35	10.83	1.33	1.87	5.53 ^c	.115	1.08
MI-A	11.88	0.35	11.50	0.84	1.03	6.35 ^c	.340	0.59

Note: HRQ_{po} = Helpful Responses Questionnaire-post, T-Eval = Training Evaluation, Tot = Total, Model = Modeling, Info = Information, Dir = Directiveness, TrainQ = Training quality, PSkills = Practice skills, TrainerQ = Trainer quality, MI-A = Motivational Interviewing adherent. Total scores were obtained for all scales.

^aSupervisor-1, *n* = 8; supervisor-2, *n* = 9.

^bSupervisor-1, *n* = 8; supervisor-2, *n* = 6.

^c*t*-test used for unequal variances.

Table 3
Differences for Students versus Research Staff on HRQ and Training Evaluation

	Research Staff		Students		Statistic			
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>df</i>	<i>p</i>	<i>d</i>
HRQ _{po} ^a	19.23	5.53	12.00	7.12	2.15	15	.048	1.13
T-Eval ^b								
Tot	126.20	5.18	112.57	12.54	2.72	7.45 ^c	.028	1.42
Model	12.00	0.01	10.86	1.77	1.71	6.00 ^c	.139	0.91
Info	11.2	1.03	9.57	2.37	1.71	7.61 ^c	.128	0.89
Dir	52.00	2.40	45.57	4.83	3.25	8.10 ^c	.011	1.69
TrainQ	15.60	0.97	13.43	1.81	2.90	8.40 ^c	.019	1.49
PSkills	11.60	0.97	10.86	1.21	1.41	15	.180	0.67
Trainer Q	11.90	0.32	10.86	1.21	2.22	6.57 ^c	.064	1.18
MI-A	11.90	0.32	11.43	0.79	1.50	7.37 ^c	.175	0.78

Note: HRQ_{po} = Helpful Responses Questionnaire-post, T-Eval = Training Evaluation, Tot = Total, Model = Modeling, Info = Information, Dir = Directiveness, TrainQ = Training quality, PSkills = Practice skills, TrainerQ = Trainer quality, MI-A = Motivational Interviewing adherent. Total scores were obtained for all scales.

^a Research staff, *n* = 13; student trainees, *n* = 4.

^b Research staff, *n* = 10; student trainees, *n* = 7.

^c *t*-test used for unequal variances.