Published in final edited form as:

J Subst Abuse Treat. 2016 September ; 68: 11–23. doi:10.1016/j.jsat.2016.04.005.

The Effectiveness and Cost of Clinical Supervision for Motivational Interviewing: A Randomized Controlled Trial

Steve Martino^{1,2}, Manuel Paris Jr.³, Luis Añez³, Charla Nich², Monica Canning-Ball², Karen Hunkele², Todd A. Olmstead^{4,5}, and Kathleen M. Carroll²

Steve Martino: steve.martino@yale.edu; Manuel Paris: manuel.paris@yale.edu; Luis Añez: luis.aneznava@yale.edu; Charla Nich: charla.nich@yale.edu; Monica Canning-Ball: monica.canning-ball@yale.edu; Karen Hunkele: karen.hunkele@yale.edu; Todd A. Olmstead: tolmstead@austin.utexas.edu; Kathleen M. Carroll: kathleen.carroll@yale.edu

¹VA Connecticut Healthcare System, Psychology Service, 950 Campbell Avenue, West Haven, CT 06516, USA

²Yale University School of Medicine, Department of Psychiatry, Division of Substance Abuse, 950 Campbell Avenue, West Haven, CT 06516, USA

³Yale University School of Medicine, The Hispanic Clinic, 25 Park Street, New Haven, CT 06511, USA

⁴The University of Texas at Austin, Lyndon B. Johnson School of Public Affairs, 2300 Red River Street, Austin, TX 78713, USA

⁵Seton/UT Clinical Research Institute, 1400 North IH 35, Austin, TX 78701, USA

Abstract

The effectiveness of a competency-based supervision approach called Motivational Interviewing Assessment: Supervisory Tools for Enhancing Proficiency (MIA: STEP) was compared to Supervision-As-Usual (SAU) for increasing clinicians' motivational interviewing (MI) adherence and competence and client retention and primary substance abstinence in a multisite Hybrid Type 2 effectiveness-implementation randomized controlled trial. Participants were 66 clinicians and 450 clients within one of eleven outpatient substance abuse programs. An independent evaluation of audio recorded supervision sessions indicated that MIA: STEP and SAU were highly and comparably discriminable across sites. While clinicians in both supervision conditions improved their MI performance, clinician supervised with MIA: STEP, compared to those in SAU, showed significantly greater increases in the competency in which they used fundamental and advanced MI strategies when using MI across seven intakes through a 16-week follow-up. There were no retention or substance use differences among the clients seen by clinicians in MIA: STEP or SAU. MIA: STEP was substantially more expensive to deliver than SAU. Innovative alternatives to resource-intensive competency-based supervision approaches such as MIA: STEP are needed to promote the implementation of evidence-based practices.

Corresponding Author: Steve Martino, Ph.D., Department of Psychiatry, VA Connecticut Healthcare Center (116B), 950 Campbell Avenue, West Haven, CT 06516, Work Phone: (203) 937-3486 x7418.

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.

Keywords

clinical supervision; motivational interviewing; substance abuse treatment; implementation strategies; cost estimation

1. Introduction

Clinical supervision is one of the most widely used strategies for teaching clinicians psychotherapy and supporting clinical implementation (Beidas & Kendall, 2010; Bernard & Goodyear, 2014; Carroll, Martino, & Rounsaville, 2010; Falender et al., 2004; Herschell, Kolko, Baumann, & Davis, et al., 2010; Martino, 2010; Rakovshik and McManus, 2010; Schoenwald, Mehta, Frazier, & Shernoff, 2013; Schoenwald, Sheidow, & Chapman, 2009; Watkins, 2011a). Through the use of distinct supervisory competencies (American Psychological Association, 2015; Borders et al., 2011; Falender & Shafranske, 2012; Bernard & Goodyear, 2014), clinical supervision has been shown to benefit clinicians by increasing their treatment knowledge, confidence, and skill (Beutler & Kendall, 1995; Holloway & Neufeldt, 1995; Wheeler & Richards, 2007). However, little empirical attention has been given to the effectiveness of clinical supervision on improving client outcomes, considered by many to be the “acid test” by which clinical supervision should be judged (Bernard & Goodyear, 2014; Ellis & Ladany, 1997; Falender & Shafranske, 2012; Lichtenberg et al., 2007). In addition, no studies have detailed the cost of supervising clinical practice, a surprising void in the literature given the widespread use of clinical supervision for training purposes. This study presents a multi-site randomized controlled trial examining the effectiveness and cost of a competency-based clinical supervision approach for motivational interviewing (MI; Miller & Rollnick, 2013), called *Motivational Interviewing Assessment: Supervisory Tools for Enhancing Proficiency* (MIA: STEP; Martino et al., 2006), on clinician MI adherence and competence and client treatment retention and outcome.

MI is a well-recognized treatment for substance use disorders that combines person-centered principles with strategies for enhancing motivation for change (Miller & Rollnick, 2013; Miller & Rose, 2009). The proficient use of MI has been shown to increase client statements that favor change (called change talk), with the balance of more pro- and less anti-change talk statements within sessions related to behavior change (Magill et al., 2014). Overall, meta-analyses have shown that MI improves client treatment retention and substance use outcomes (Burke, Arkowitz, & , 2003; Hettema, Steele, & Miller, 2005; Lundahl, Kunz, Brownell, Tollefson, & Burke, 2010; Smedslund, Berg, Hammerstrom, Steiro, Leiknes, Dahl, & Karlsen, 2011). A competency-based supervision approach has been touted as a promising strategy for promoting proficient MI practice (de Roten, Zimmermann, Ortega, & Delpland, 2013; Madson, Loignon, & Lane, 2009; Schwalbe, Hans, & Zweben, 2014).

Competency-based clinical supervision is an approach that explicitly identifies the knowledge and skills that clinicians need to deliver psychotherapy appropriate to their clinical settings and clientele (Falender & Shafranske, 2007). Core elements of high quality competency-based supervision include (a) directly observing clinicians' practice in sessions

or reviewing audio or video recorded ones, (b) using performance feedback to monitor practice, and (c) providing individualized coaching to further develop clinicians' knowledge and skills (APA, 2015; Falender & Shafranske, 2012; Reisner & Milne, 2012; Watkins & Scaturro, 2013) – which parallels the elements used to supervise clinicians in psychotherapy efficacy and effectiveness trials (Baer et al., 2007; Beidas & Kendall, 2010; Carroll, Martino, & Rounsaville, 2010; Hershell et al., 2010). The competency-based supervision approach has been adopted for use across multiple psychotherapeutic theoretical perspectives (Farber & Kaslow, 2010) within the United States (APA, 2015; Borders et al., 2011) and internationally (Gonsalvez & Milne, 2010). Training professionals in competency-based clinical supervision is now recognized as a critical area of workforce development (Center for Substance Abuse Treatment, 2007; Fleming, 2004; Hoge, Migdole, Farkas, Ponce, & Hunnicutt, 2011; Miller, Sorensen, Selzer, & Brigham, 2006; Roche, Todd, & O'Connor, 2007; Watkins, 2011b).

In 2001, the National Institute on Drug Abuse (NIDA) and the Substance Abuse and Mental Health Services Administration Center for Substance Abuse Treatment Addiction Technology Transfer Centers (SAMHSA/CSAT/ATTC) collaborated to develop training products that would support the dissemination and implementation of research findings from NIDA-funded treatment studies into community-based practice (Condon, Miner, Balmer, & Pintello, 2008). One product, MIA: STEP (Martino et al., 2006), was developed to support local program-based supervision of MI. MIA: STEP adapted the supervision methods used to train clinical supervisors in several MI effectiveness trials (Ball et al., 2007; Carroll et al., 2006; Carroll et al., 2009), consistent with the competency-based clinical supervision approach. It aims to improve clinicians' adherence and competence using MI following initial workshop training, thereby contributing to better client treatment retention and outcome. Interest in MIA: STEP has been strong, and a cadre of national trainers has been prepared to provide MIA: STEP training in the United States via the SAMHSA/CSAT/ATTC network (Martino et al., 2010). To date, the effectiveness and cost of using MIA: STEP in community treatment programs has not been determined.

Multiple reviews of supervision research have been conducted (Ellis & Ladany, 1997; Ellis, Ladany, Krenzel, & Schult, 1996; Freitas, 2002; Holloway & Neufeldt, 1995; Milne & James, 2000; Watkins, 2011a; Wheeler & Richards, 2007). Broadly, these reviews suggest that clinical supervision enhances clinicians' treatment knowledge, adherence, competence, self-confidence, and clinician-client relationship (e.g., therapeutic alliance, satisfaction). Further, clinical supervision has been associated with a reduction in clinicians' emotional exhaustion and intention to quit their jobs (Knudsen, Ducharme, & Roman, 2008). Active training techniques (e.g., performance feedback, coaching via behavioral rehearsal/role-play) employed in successive supervision sessions maintain and sometimes additionally improve clinicians' therapeutic skills following initial didactic or workshop training (Beidas & Kendall, 2010; Hershell et al., 2010; Milne, Sheikh, Pattison, & Wilkinson, 2011; Rakovshik & McManus, 2010). These findings apply across a range of psychotherapies, including cognitive behavioral therapy (Mannix et al., 2006; Sholomskas et al., 2005), multisystemic therapy (Schoenwald et al., 2009), problem-solving therapy (Bambling, King, Raue, Schweitzer, & Lambert, 2006), and MI (de Roten, Zimmermann, Ortega, & Delpland, 2013; Madson, Loignon, & Lane, 2009; Martino, et al., 2010; Miller et al., 2004; Schwalbe, Hans,

& Zweben, 2014; Smith et al., 2012; Soderlund, Madson, Rubak, & Nilsen, 2011). A recent meta-analysis of MI training studies showed that the addition of approximately monthly post-workshop supervisory feedback and coaching sessions over a 6-month period was sufficient to sustain workshop training effects, with an overall MI skills training effect size of .75 (Schwalbe et al., 2014).

In contrast to the numerous studies examining the effectiveness of clinical supervision on clinicians' knowledge, skills, and attitudes, high quality research examining the impact of clinical supervision on client outcomes is scant. Only 18 supervision-client outcome studies were conducted from 1981 to 2006 (Watkins, 2011a), and the vast majority of these studies were marked by several methodological shortcomings. Shortcomings included 1) lack of manuals operationalizing supervision, 2) insufficient documentation of supervisor training, 3) failure to demonstrate supervision integrity, 4) poorly defined and psychometrically weak instruments to measure clinician and client outcomes, 5) very small sample sizes of supervisors and clinicians reducing power for hypothesis testing, 6) inclusion of clinicians unrepresentative of community program service providers, and 7) no comparison or control conditions, random assignment, or routine follow-up assessments (Ellis & Ladany, 1997; Ellis et al., 1996; Herschell et al., 2010; Holloway & Neufeldt, 1995; Reisner & Milne, 2014; Roth, Pilling, & Turner, 2010; Watkins, 2011a; Schoenwald et al., 2009). Moreover, none of the studies controlled for the effect of seminars or workshops that preceded supervision.

An area absent in supervision research is the cost of clinical supervision. Competency-based clinical supervision approaches require a significant expenditure of time and effort to train supervisors, observe clinician practice, gather performance feedback data, and conduct the supervision sessions. The costs associated with these expenditures are relatively unknown, though it is presumed to be one of the single most expensive investments in supporting the implementation of effective treatments (Gonsalvez & Milne, 2010). In one study, Olmstead and colleagues (2012) estimated that the cost of training two program-based supervisors to clinically supervise eight clinicians in MI monthly over a three-month period was approximately \$9,700.00 (in 2006 US dollars). The cost of providing supervision needs to be justified by the capacity of competency-based supervision to achieve significantly better clinician treatment integrity and client outcomes than standard supervisory practices used in community treatment programs. Cost and efficacy data could help influence behavioral health care policy, including decisions about the dissemination and fiscal support state and federal agencies provide for clinical supervision.

This study presents a multi-site randomized controlled trial testing the effectiveness of MIA: STEP supervision, compared to supervision-as-usual (SAU), in improving community treatment program clinicians' MI adherence and competence within client intake sessions, as well as the program retention (percent scheduled sessions attended, percent retained in treatment) and days of primary substance abstinence of clients receiving MI in the study. We hypothesized that MIA: STEP would result in significantly better MI adherence and competence, better client retention, and more days of primary substance abstinence than SAU. We also calculated the cost of providing supervision in both arms.

2. Materials and methods

2.1. Study design and overview

This study uses a Hybrid Type 2 effectiveness-implementation randomized controlled trial design (Curran, Bauer, Mittman, Pyne, & Stetler, 2012) in that it simultaneously tests MIA: STEP, an implementation strategy, and MI, the clinical intervention being implemented. Supervisors, clinicians, and clients within 11 outpatient community treatment programs participated in the study. Based on a power analysis to account for the multi-level and multi-site nature of the study (Raudenbush, 1997; Raudenbush & Xiaofeng, 2003), we targeted enrollment at 60 clinicians (4 or 6 per site) and 420 clients (28 or 42 per site) across sites. The study protocol and informed consent procedures were reviewed and approved by Yale University School of Medicine Human Investigation Committee. A Data Safety and Monitoring Board convened for this study reviewed serious adverse events ($n = 27$, none of which were study related nor different between the conditions) for the duration of the protocol. Recruitment of participants occurred over a 50-month period from April 2009 to June 2013.

2.2. Setting

The study was conducted in 11 not-for-profit, non-opiate agonist outpatient community addiction treatment programs based in Connecticut, United States. Although the programs had an array of available treatment options (individual, group, couples/family, rehabilitative services), clients in these programs typically received group treatment on a weekly or more intensive several time per week schedule. All programs required clients to be clinically evaluated within a single intake session as a prelude to the start of treatment. Clinician and client characteristics and length of intake sessions and treatment varied across programs (see below).

2.3. Participants

2.3.1. Clinicians—To be eligible, clinicians had to agree to work at least 20 hours per week, expect to be employed at the program during the trial, and agree to have MI intakes and supervision sessions recorded. Clinicians could not have had past MI supervision involving direct observation with rating-based feedback or coaching. A total of 76 clinicians were screened for interest and eligibility; of these, eight dropped out prior to randomization following workshop training (see Figure 1, CONSORT diagram) for varying reasons (left agency, medical leave, administrative disapproval, decided not to participate). Two clinicians dropped out after randomization but before receiving supervision and seeing client participants. Therefore, 66 clinicians participated in the study. These clinicians were predominantly female (79%), Caucasian (65%; 20% Hispanic; 14% African American; 1% multiracial) and had a mean age of 41.3 ($SD=13.2$) years. Most (68%) had master's degrees, and the remainder was 14% bachelor's, 8% associate's, 9% high school, and 1% doctorate degrees. About 44% had received state certification as a substance abuse counselor; 52% were licensed in their fields. Clinicians averaged 8.9 ($SD=8.7$) years of general and 7.7 ($SD=8.8$) years of substance abuse counseling experience. Most (77%) had read Miller and Rollnick's (2002) textbook within a year prior to participation, but few had past year MI training. Clinicians were similar across conditions (see Table 1). Statistically significant site

differences indicated two sites had older clinicians ($F(1,10) = 4.54, p < .000$) with more years of general counseling ($F(1,10) = 5.62, p < .000$) and substance abuse counseling ($F(1,10) = 5.88, p < .000$) experience.

2.3.2. Clients—Client participants were non-opiate agonist outpatient treatment seekers who had used their primary substance in the past four weeks and were 18 years of age or older, English-speaking, and willing to have their intake sessions audio recorded. Clients were ineligible if they were too medically or psychiatrically unstable to participate in outpatient treatment, highly unlikely to be reached for follow-up due to residential instability or imminent incarceration, or seeking detoxification only, methadone maintenance or buprenorphine treatment, or residential inpatient treatment. A total of 1520 clients were screened for interest and eligibility to participate in the study, with 450 being randomized and 1070 ineligible for randomization (see Figure 2, CONSORT diagram) mainly because of an absence of self-reported primary substance use in the past four weeks or disinterest in delaying their intake in order to complete the study's baseline assessments.

The 450 randomized clients were on average 35 (SD=11.6) years old, mostly single (61%), unemployed (67%), and ethnically diverse (46% African American; 28% Caucasian; 21% Hispanic; 4% multiracial; and 1% other). All participants were current substance users, with many meeting current criteria for DSM-IV substance use disorders: alcohol (39%), cannabis (35%), cocaine (28%), opioids (7%), hallucinogens (3%), and stimulants (1%), based on the Structured Clinical Interview for DSM-IV (Alcohol and Drug Use Disorders sections only) (Kransler, Kadden, Babor, Tennen, & Rounsaville, 1996). Clients designated their primary substance problem (i.e., self-identified as causing the most problems) as: alcohol (32%), cannabis (34%), cocaine (24%), opioids 6%, and other (4%), with most having used alcohol (80%) and cannabis (57%) four weeks prior to randomization, including a large minority who had used cocaine (38%). On average, clients had been abstinent from their primary substance for about 14–15 days. About half the clients were court-mandated to treatment. In general, clients presented to treatment with self-reported recognition of their alcohol or drug problems, openness to reflect about the pros and cons of their substance use, and already having taken some steps to change their drinking or drug use, as evidenced by their moderately high scores on the Stages of Change Readiness and Treatment Eagerness Scale (SOCRATES) (Miller & Tonigan, 1996). Client characteristics were similar between conditions (see Table 2). However, there were several statistically significant site differences, including gender ($X^2(1) = 30.03, p = .001$; proportion of female clients ranged from 13% to 54%), ethnicity ($X^2(1) = 96.54.03, p < .000$; proportions ranged as follows: 10–70% Caucasian; 7–68% African-American; 8–33% Hispanic), employment ($X^2(1) = 42.01, p = .003$; proportion of unemployed clients ranged from 40% to 83%), treatment mandated status ($X^2(1) = 100.82, p < .001$; proportion of clients mandated to treatment ranged from 15% to 80%), and primary substance use ($X^2(1) = 142.20, p < .001$). Regarding the latter, marijuana was the most frequently reported primary substance at five sites (range = 40% to 58%), alcohol at two sites (47%, 63%), and cocaine at two sites (46%, 47%). Two sites had equal proportions of primary marijuana, alcohol, and cocaine users.

2.4. Supervision selection and procedures

Two supervisors from each program participated in the study ($n=22$), one for each of the respective conditions. MIA: STEP and SAU supervisors had to work a minimum of 20 hours per week at the program, expect to be employed at the program during the trial, and agree to have supervision audio recorded. MIA: STEP supervisors had the additional requirements of having past MI seminar or workshop training and receiving program administration approval to provide MIA: STEP supervision, given the added time burden. Research staff worked with program administrative leadership to identify potential study supervisors, describe the study to them, and obtain their consent to participate. A combination of MI training experience and preference typically dictated a supervisor's assignment to MIA: STEP. The 22 supervisors were demographically similar to the clinicians, except that they had higher percentages of earning a master's degree (82%), alcohol and drug counseling certification (64%), professional licensure (96%), and average years of general counseling experience (17.4, $SD=9.2$). There were no statistically significant differences in supervisor demographic characteristics across conditions.

2.4.1. Clinician recruitment—The study was presented at staff meetings at each program. Clinicians interested in participation subsequently met with a research assistant, who then screened clinicians for eligibility and obtained written informed consent and baseline assessments. Clinicians received \$50 per completed assessment point.

2.4.2. Workshop training for MI intake—MIA: STEP supervisors and all clinicians within each program received an eight-hour, on-site MI workshop conducted by the first author per Motivational Interviewing Network of Trainers recommendations (Miller & Rollnick, 2013). Training focused on how to apply MI in an intake. All participants received an individual copy of a MI textbook (Miller & Rollnick, 2002) and a detailed MI intake manual used in our previous trial (Carroll et al., 2006).

2.4.3. MIA: STEP training and consultation—Following workshop training, MIA: STEP supervisors completed and audio recorded a MI intake with a client who met study eligibility criteria. These sessions were reviewed by one of two MIA: STEP experts (MP, LA). The supervisor then received MIA: STEP supervision on the case from the expert to directly experience MIA: STEP and have it modeled for them. Supervisors also completed a 15-hour MIA: STEP workshop conducted by the first author and one of the experts (MP), consistent with the published MIA: STEP training-of-trainers curriculum (Martino et al., 2006). In addition, each supervisor had monthly on-site or phone consultations throughout the trial from the MIA: STEP expert based on review of co-rated MI intakes and the associated recorded supervision session. Consultation provided a means to support MIA: STEP integrity throughout the trial.

2.4.4. Clinician randomization and trial initiation—Given that clinicians variably increase MI skills from workshop training (Martino, Canning-Ball, Carroll, & Rounsaville, 2011; Schwalbe et al., 2014), following the workshop each clinician completed and audio recorded a MI intake with a client who met study eligibility criteria, and these sessions were independently rated for adherence and competence using the *Independent Tape Rater*

System (ITRS; Ball, Martino, Corvino, Morgenstern, & Carroll, 2002). Based on these session ratings, clinicians were stratified by performance into blocks of two and then randomized (coin flip) to condition to reduce risk of imbalance in pre-supervision MI skills. After randomization, clinicians assigned to MIA: STEP received supervision on their post-workshop MI intake from their MIA: STEP supervisors, who had separately listened to and rated them, to ensure clinicians had received an initial “dose” of MIA: STEP supervision before they saw their first client participants in the trial. Clinicians in SAU were free to receive supervision from the SAU supervisor after the workshop as it naturally occurred within the program.

2.4.5. Supervision conditions—MIA: STEP (Martino et al., 2006) includes a system for rating 10 MI consistent items for adherence and competence and five MI inconsistent items for adherence, all derived from the ITRS (Ball et al., 2002), and tools for providing rating-based feedback and coaching (teaching tools, descriptions of MI consistent strategies, guidelines and a format for MI consistent supervision, audio recorded, transcribed and rated MI intake sessions). After a clinician completed a MI intake, the MIA: STEP supervisor listened to and rated it, completed the feedback form, and met with the clinician to discuss the session and coach skill development before another client was seen. MIA: STEP supervisors received \$30 per completed supervision session. All MIA: STEP supervision sessions were audio recorded for monitoring and independent integrity assessment. MIA: STEP supervision was conducted up to seven times per clinician. Clinicians did not see their next client in the protocol until they had received MIA: STEP supervision for their prior case.

SAU consisted of the practices typically used at each program to supervise clinicians’ intakes. Supervisors and clinicians were instructed to meet as they typically would to supervise intake sessions, and new intake assignments were not contingent on SAU supervision having occurred on the prior case. These sessions were not recorded to minimize alteration of SAU practices. However, clinicians were asked to complete a checklist when supervision of study-related client intakes occurred to characterize what happened in SAU. In addition, on an approximately monthly basis, research staff asked clinicians if they had met with their SAU supervisors and had them complete the checklist if SAU had occurred. Moreover, to independently identify SAU practices and the extent to which SAU discriminated from MIA: STEP, before workshop training and random assignment had occurred, the two program supervisors supervised one of each clinician’s intake sessions in their usual manner, splitting the clinicians between them based on common schedules. At the end of the trial, only the SAU supervisors conducted this task again with the clinicians who were assigned to SAU. These baseline and post-trial supervision sessions were recorded and independently rated (see below).

2.5. MI Intervention procedures

2.5.1. Client recruitment and randomization—Research staff integrated client recruitment into each site’s admissions processes, which included calling clients before scheduled intake appointments or approaching them on scheduled intake days. Eligible clients completed written informed consent and baseline assessments and were randomized

to condition with a computerized urn randomization program (Stout, Wirtz, Carbonari, & Delboca, 1994) to ensure balance within sites and across conditions on the primary substance used (alcohol/drug) and whether the client had been legally mandated to treatment (yes/no). Per randomization, clients were scheduled for MI intakes with either a MIA: STEP or SAU clinician within two weeks of randomization. Across sites, the mean number of days elapsing from randomization to the intake was 5.0 ($sd = 4.3$). Clinician availability within condition dictated who conducted the intake. Clients received \$30, \$50, and \$75 for completing the respective baseline, 4-week and 12-week follow-up assessments.

2.5.2. MI Intake—MI was used in an intake as a means to improve client retention, consistent with our previous MI intake protocol (Carroll et al., 2006), and because instituting a multi-session individual MI intervention was incompatible with the largely group treatment settings. The MI intake was construed as a “MI sandwich” (Martino et al., 2006) in that (1) the first 20–30 minutes of the intake focused on engaging clients, understanding their primary substance use, and enhancing their motivation for treatment as a means for cutting back or quitting substance use, and (2) the last 10–15 minutes focused on making plans for change, if ready, or alternative actions if not ready to enter treatment or change substance use. A standard psychosocial intake assessment (30–60 minutes) was sandwiched in between the MI parts, and clinicians were trained on how to transition from one part to the next. Both MIA: STEP and SAU clinicians were asked to conduct MI intakes with all client participants. All MI intakes were audio recorded for MI adherence and competence assessment. On average MI intakes were 1.3 ($SD=.4$) hours in length, with no significant differences between supervision conditions. Significant site differences ($F(1,10) = 3.9, p < .001$) showed one site had significantly longer intake sessions, which lasted on average 2.0 ($SD=.4$) hours.

2.6. Assessment procedures

Clinicians completed assessments at baseline, after completing 7 MI intakes (post-trial), and 16-weeks post-trial. Clients completed assessments at baseline, 4-week, and 12-week follow-up. In addition, supervision integrity (i.e., supervisor adherence and competence) and cost were independently assessed at the end of the study.

2.6.1. Clinician assessments—To measure change in MI adherence and competence over time, in addition to trial sessions, each clinician recorded a baseline intake session in which the clinician was instructed to motivate the client (who met study eligibility criteria) to change his or her substance use; this same procedure was used at the 16-week post-trial follow-up. Primary clinician outcomes were measured by the Independent Tape Rater Scale (ITRS).

The ITRS (Ball, Martino, Corvino, Morgenstern, & Carroll, 2002) is a reliable and valid measure (Martino et al., 2008; Santa Ana et al., 2009; Gibbons et al., 2010) that assesses adherence and competence in MI and strategies inconsistent with MI. It has been used in several large multi-site effectiveness trials (Ball et al., 2007; Carroll et al., 2006; 2009) and clinician training research studies (Martino et al., 2010; Martino et al., 2011). Raters use a 7-point Likert-type scale to reflect the frequency or extensiveness of each MI strategy

(adherence: 1 = not at all, to 7 = extensively) and the skill or competence with which the strategy is deployed (competence: 1 = very poor, to 7 = excellent). A prior confirmatory factor analysis of ITRS supported a two-factor solution (Martino et al., 2008) for the sum total of MI consistent items, representing the mean scores for five fundamental MI strategies that tap the client-centered aspects and basic skills of MI (e.g., reflection, fostering collaboration) and five advanced MI strategies related to evoking and mobilizing motivations for change (e.g., developing discrepancy between substance use and goals, change planning). The mean adherence score for five MI-inconsistent items (e.g., unsolicited advice, direct confrontation) is also calculated.

Ten raters were trained to rate the audio recorded sessions using the ITRS, according to procedures described by Martino et al. (2008). Raters (blinded to study training condition, assessment point, and site) attended a seminar to learn how to rate ITRS items and then rated an identical set of 12 calibration recordings selected randomly from the larger pool of recorded intakes, which were used to evaluate ITRS item inter-rater reliability. Each recording was time stamped to capture the first 30 minutes and last 15 minutes of each session, namely, the segments of the session where the majority of MI would likely occur using the “MI sandwich” intake approach. Raters only rated these segments. According to rating criteria established by Cicchetti (1994), two-way mixed effects (fixed effects for raters and random effects for sessions), average-measures, intra-class correlation coefficients (ICC) for the fundamental and advanced MI strategy adherence and competence scores showed good to excellent inter-rater reliability (adherence ICC: fundamental = .89 and advanced = .92; competence ICC: fundamental = .70 and advanced = .86). Inter-rater reliability for the MI-inconsistent adherence items was excellent (ICC = .91). Raters were assigned MI intake recordings randomly. Each recording was rated by only one rater.

2.6.2. Client assessments

2.6.2.1. Treatment retention: Because client treatment between sites varied in frequency (1–4 days per week), intensity (1–3 sessions per day), and length (4 weeks or longer), for each client we calculated the cumulative total number of sessions attended and divided this sum by the total number of sessions scheduled to attend to obtain the percentage of sessions attended as a common metric of treatment retention across sites. We also calculated the percentage of clients who remained enrolled in treatment at each follow-up point. Client attendance and enrollment data were extracted from administrative records.

2.6.2.2. Substance use: The Substance Use Calendar (SUC) is an interview assessment of self-reported substance use (marijuana, cocaine, alcohol, methamphetamine, benzodiazepines, opioids, and other drugs) completed at each contact by a research assistant. Adapted from the Time Line Follow-Back interview (Sobell & Sobell, 1992), a reliable and valid instrument for monitoring substance use (Fals-Stewart, O’Farrell, Freitas, McFarlin, & Rutigliano, 2000), the SUC uses a similar calendar method to allow for a continuous evaluation of daily substance use (Carroll et al., 2004).

Breathalyzers (2009–2011 AlcoHAWK Precision Elite, 2011–2013 BACtrack S75 Pro Series) and OnTrack Testcups (5.01; Roche Diagnostics, Basel, Switzerland) were used at all

programs for breath and urine screening. Cut-off scores for urine drug screens were: amphetamines (100 ng/ml), cocaine (300 ng/ml), methamphetamine (500 ng/ml), morphine (300 ng/ml), benzodiazepines (20 ng/ml), and THC (50 ng/ml). Analyses of baseline breath and urine samples indicated high correspondence with clients' self-reported recent substance use. Only 25 breathalyzer samples had readings above .08 (3%). Rates of positive drug urine test results when the clients had denied recent use were low and were comparable to our previous MI intake trial (Carroll et al., 2006): cocaine/stimulants (6.9%); marijuana (11.5%); opiates (2.7%), and benzodiazepines (6.5%).

2.6.2.3. Treatment utilization: The Treatment Utilization Form (TUF) is an 11-item interview that was adapted and shortened from the Treatment Services Review (McLellan, Alterman, Cacciola, Metzger, & O'Brien, 1992) to measure the extent of additional program services (e.g., non-study related individual, group, and family counseling; 12 step meetings; medical or vocational appointments) received either at the program or by referral to another community service.

2.6.3. Supervision integrity assessment—Supervision integrity was primarily assessed using the Supervisor Adherence and Competence Scale (SACS). Adapted from the Yale Adherence and Competence Rating System (Carroll et al., 2000), SACS (Martino, Paris, Añez, & Carroll, 2012) is a 12-item independent rating scale that assesses the frequency (adherence: 1 = not at all, to 7 = extensively) and skill (competence: 1 = very poor, to 7 = excellent) in which the supervisors use five MIA: STEP-consistent (e.g., fostering collaboration, feedback, coaching) and six general supervision strategies (e.g., case presentation and management, treatment planning), as well as a 7-point Likert-type general rating to assess the degree to which supervision was delivered in a style consistent with MI (e.g., collaborative, affirming, supervisee-centered). Four mean item scores are calculated for MIA: STEP and general supervision adherence and competence items.

Using the same procedures to prepare ITRS raters, five raters were trained to rate the audio recorded supervision sessions using the SACS. All raters were blinded to study training condition, assessment point, and site. After initial seminar training, raters rated an identical set of eight calibration recordings (4 MIA: STEP, 4 SAU) selected randomly from the larger pool of recorded supervision sessions. Inter-rater reliability for MIA: STEP consistent adherence (.96) and competence (.95) and general supervision adherence (.93) and competence (.89) was excellent. The ICC for MI supervision style was .96. Raters were assigned supervision session recordings randomly. Each recording was rated by only one rater.

In addition, a 3-item supervision checklist was completed by SAU clinicians when they received supervision for any study MI intake. The items ask about: 1) type of supervision received (individual, group); 2) areas covered (assessment, treatment planning, administrative case review); and 3) supervisory techniques (e.g., provision of education/information, general case discussion, feedback, coaching).

2.6.4. Cost estimation—Cost data for supervision was collected prospectively in the trial by research staff. A research assistant administered a survey to the Chief Financial

Officer, Chief Executive Officer, or accountants at each site to obtain necessary cost information (e.g., program overhead and fringe rates; unit cost of space). In addition, expert trainers, supervisors, and clinicians reported their annual salary/hourly wage, and research staff collected data on resources used (e.g., space utilized for activities, number of supervision sessions held and time spent in them by supervisors and clinicians, time spent by expert trainers/consultants in activities, time spent by MIA: STEP supervisors rating MI intakes and preparing feedback). Costs were calculated from the perspective of the program and adjusted to 2014 US dollars using the Consumer Price Index. All labor costs included fringe benefits and overhead. Research-specific costs (e.g., incentive payments for study participation) were excluded from the analysis. Major cost categories used to derive the total cost of MIA: STEP and SAU supervision included 1) MI and MIA: STEP workshops, 2) post-workshop practice cases, 3) MIA: STEP expert consultations, 4) supervision, and 5) self-study activities.

2.7. Data analysis

We used SPSS-21 (IBM SPSS Statistics for Windows, Version 21.0) for all data analyses. We first tested for discrimination between the two supervision conditions for MIA: STEP consistent and general supervision strategy adherence and competence and the degree to which the supervisors adopted a MI style of supervision via Supervision Condition X Site ANOVAs.

2.7.1. Clinician outcomes—Separate hierarchical linear regression models were completed for the five primary outcomes (fundamental and advanced MI strategy adherence and competence, MI inconsistent strategy adherence) from baseline to the end of the supervision trial, and baseline to 16-weeks post-trial, with supervision condition as the independent condition and site nested within supervision condition as a random factor. All outcome analyses were conducted with the intent-to-treat ($n = 450$) sample. We used Feingold's (2009) d formula for estimating effect magnitude in hierarchical linear regression models.

2.7.2. Client outcomes—For the primary outcome (percent sessions attended) and one secondary outcome (days primary substance abstinence), we used separate regression models, with supervision condition as a fixed factor and site as a random factor for both the 4-week and 12-week follow-ups. For these models, partial eta squared provided an estimate of the variance accounted for by supervision condition. For the dichotomous outcome (whether the client was still retained in treatment at the 4- and 12-week follow-ups), we used a binomial logistic regression model and reported Wald chi-squares. All outcome analyses were conducted with the intent-to-treat ($n = 450$) sample.

2.7.3. Multiple-comparison correction—To control for multiple comparisons, we used a familywise alpha of $p < .01$ (.05/5) for the evaluation of the supervision integrity and the five clinician outcomes, and $p < .017$ (.05/3) for the three client outcomes.

2.7.4. Cost—For each supervisor condition, cost estimates are reported in total and by major cost category.

3. Results

3.1. Retention

Of the 66 randomized clinicians, 66 (100%) conducted at least one MI intake, 50 (76%) completed the post-trial assessment, and 39 (59%) completed the 16-week post-trial assessment. Reasons for dropout typically involved leaving the site (e.g., moved, obtained another job, terminated), unexpected medical or family leave, or being too busy at work to continue study participation. Rates of follow-up did not differ by condition, but did vary across sites. Two sites had <50% 16-week clinician retention rates because several clinicians had left the site. At another site, slow client recruitment disabled the staff from completing the trial before the end of the grant's funding period and, thus, no post-trial or 16-week post-trial assessments were collected there, though each clinician conducted two to four MI intakes.

Of the 450 randomized clients, 411 (91%) completed a MI intake, 408 (91%) completed the 4-week follow-up, and 385 (86%) completed the 12-week follow-up. Rates of follow-up did not differ by condition or by site.

3.2. Supervision delivery and integrity

During the trial, a total of 247 MIA: STEP supervisions occurred, with each clinician receiving on average 6.5 (SD = 3.3) sessions. Mean time spent in MIA: STEP supervision was 35 minutes (SD = .13.2). In comparison, only five individual SAU supervisions occurred during the trial, all occurring at one site with two SAU clinicians and each 20 minutes or less in duration. SAU sessions focused on general case review for the purposes of diagnostic assessment and treatment planning. Clinicians remained in the trial on average 7.3 months (SD = 3.8). At one site, clinicians remained in the trial significantly longer ($F(1,10) = 5.12, p < .000; M = 13.4, SD = 5.11$) because of slow client recruitment.

To determine if MIA: STEP was discriminable from SAU practices across sites, a random selection of 155 of 247 (63%) audio recorded MIA: STEP and 55 audio recorded baseline and post-trial SAU supervision sessions were independently rated for supervision integrity. Results show (see Table 3) that there were significant effects for Supervision Condition and Site, but no significant Condition X Site interactions for MIA: STEP consistent and general supervision strategies or MI style of supervision delivery, indicating that MIA: STEP was consistently discriminable from SAU supervision across the eleven sites. Overall, MIA: STEP in comparison to SAU supervisors had significantly higher MIA: STEP consistent strategy adherence and competence, lower general supervision strategy adherence delivered with similar skill, and more extensive demonstrations of a MI style of supervision in sessions. Significant site main effects showed that supervision provided at one site, compared to two other sites, was consistently higher in MIA: STEP adherence ($M = 5.2, SD = 1.2$ vs. $M = 2.8, SD = 1.6; M = 3.2, SD = 2.1$) and competence ($M = 5.5, SD = 1.2$ vs. $M = 3.5, SD = 1.5; M = 3.7, SD = 1.6$) and MI supervision style ($M = 4.6, SD = 1.5$ vs. $M = 2.2, SD = 1.6; M = 2.3, SD = 1.4$), with the remaining sites falling in between on these dimensions. General supervision adherence and competence did not significantly differ across sites.

3.3. MIA: STEP versus SAU effectiveness for clinician MI adherence and competence

Table 4 lists outcome variables by supervision condition and trial case. The results show that over the course of the trial, clinicians significantly improved their fundamental and advanced MI adherence and competence. No differential supervision effects were present for fundamental (somewhat to quite a bit present) or advanced (infrequently to somewhat present) MI adherence over the trial. However, MIA: STEP compared to SAU clinicians showed significantly greater improvement in fundamental MI competence from baseline to 16-week follow-up. In addition, MIA: STEP clinicians, compared to those in SAU, showed significantly greater improvement in advanced MI strategy competence from baseline to the end of the trial and through the 16-week follow-up. Throughout the trial, clinicians in both conditions rarely used MI-inconsistent strategies.

3.4. MIA: STEP versus SAU effectiveness for client treatment retention and primary substance use outcomes

Between conditions, as measured by the TUF, clients attended the same average number of post-intake individual ($M = .35$, $SD = 1.59$) and group sessions ($M = 1.26$, $SD = 6.35$), psychiatric medication consultations ($M = .23$, $SD = .97$), self-help groups ($M = 1.52$, $SD = 4.88$), childcare ($M = .40$, $SD = 2.73$) and family services ($M = .12$, $SD = .78$), and vocational ($M = .67$, $SD = 2.89$) and legal services ($M = .41$, $SD = 1.50$). Therefore, no co-variation of TUF variables were included in the treatment retention and primary substance abstinence outcome analyses.

3.4.1. Treatment retention—Table 5 lists outcomes variables by supervision condition. There were no significant differences between the supervision conditions in the percentage of scheduled sessions clients attended, nor in the overall percentage of clients who were retained in treatment at the 4-week and 12-week follow-ups. At four weeks, clients attended about 60% of their scheduled sessions with slightly more than 70% remaining in treatment. At 12 weeks, clients similarly attended about 60% of their scheduled sessions, with about 50% still in treatment.

3.4.2. Primary substance abstinence—Table 5 shows that there were no differences between supervision conditions for the total number of days on which a client reported remaining abstinent from his or her identified primary substance of use. At both the 4-week and 12-week follow-ups, clients reported not using their primary substances for an average of 22 of the past 28 days, compared with their baseline 14–15 days of abstinence (about a 50% increase). Subgroup analyses for respective groups of primary alcohol, marijuana, and cocaine users, the three largest subgroups in this study, also failed to demonstrate significant condition differences.

3.5. Association between clinician and client outcomes

There were no significant associations (all $ps > .05$) between the clinician's mean fundamental or advanced MI adherence and competence scores and the clients' percentage of sessions attended or days of primary substance abstinence at the 4- and 12-week follow-ups. Likewise, the clinicians' mean MI inconsistent adherence was not significantly related to client outcomes (all Pearson product-moment correlation coefficients fell within $-.10$ to

-.04 range). Finally, Supervision Condition X Retention Status ANOVAs showed no significant differences (all p s > .05) in the clinicians' mean fundamental and advanced MI adherence and competence and mean MI inconsistent adherence between those clients retained and not retained in treatment in either condition at the 4- and 12-week follow-ups.

3.6. Comparison of supervision costs

Table 6 compares supervision costs in total and by major cost category. For MIA: STEP, the two costliest categories were consultations between experts and supervisors (\$56,077) and supervisions between supervisors and clinicians (\$40,676), whereas for SAU the two costliest categories were the MI workshops (\$18,236) and time spent by clinicians self-studying MI materials (\$9,680). Across both supervision conditions, the largest component of total cost was labor (85.7%), followed by travel (9.8%), materials (3.2%), miscellaneous (0.7%), and space (0.5%).

4. Discussion

We demonstrated that community program clinicians gradually increased their fundamental and advanced MI adherence and competence over a trial of seven supervised post-workshop intake sessions in which clinicians used MI, as well as 16 weeks post-trial. Clinicians supervised with MIA: STEP, however, increased the competence in which they used fundamental and advanced MI strategies significantly more than clinicians who were supervised as usual. Supervision conditions showed no differential effects for increased MI adherence. Conversely, this study did not demonstrate differences in short-term client retention or outcomes. Clients who received intakes from clinicians supervised by MIA: STEP or as-usual had similar rates of scheduled sessions attended, program retention, and days of primary substance abstinence at the 4-week and 12-week follow-ups. Finally, we estimated that MIA: STEP was substantially more expensive to deliver compared to usual supervisory practices.

A critical component in testing the effectiveness of supervision is verifying that the supervision was delivered as intended (Milne, 2014). Analyses based on independent ratings of the audio recorded supervision sessions suggested that across sites MIA: STEP was highly discriminable from SAU. Supervisors who delivered MIA: STEP used core competency-based supervision components, which included performance-based feedback and skills coaching. In addition, MIA: STEP strategies were delivered significantly more competently and in a style more consistent with MI in MIA: STEP than SAU. These results suggest that program-based supervisors can learn to deliver competency-based supervision approaches such as MIA: STEP effectively, at least when provided with the types of supports used in this study (e.g., 2-day workshop training, monthly MIA: STEP expert consultations, payment). Moreover, clinicians assigned to MIA: STEP were supervised on average 6.5 times for half an hour per contact over a period of seven months - a 'dose' consistent with what prior MI training studies have found related to sustaining MI training effects over time (Schwalbe et al., 2014).

This study showed that clinicians supervised with MIA: STEP, more than those in SAU, significantly increased the competency in which they used advanced MI strategies at the end

of the trial and fundamental and advanced MI strategies at the 16-week follow-up point. Clinicians in both supervision conditions demonstrated similar increases in the frequency in which they used fundamental and advanced MI strategies (i.e., adherence) at both assessment points. Across conditions, MI inconsistent strategies were virtually absent. Thus, the main effect of MIA: STEP appears to be on enhancing the clinicians' MI competence, not adherence, following a basic MI skill-building workshop. Post-workshop, clinicians who received individualized feedback and coaching via MIA: STEP moved from somewhat less than adequate performance to adequate to good skill levels, whereas those in SAU continued to deliver MI at or below adequate levels of competency. This effect is strongest for advanced MI strategies post-trial ($d = .84$) and 16 weeks after MIA: STEP ended ($d = .82$). One explanation for these findings may be that the MI workshop, self-training materials (MI book and treatment manual), and repeated application of a manualized version of MI (the MI sandwich) were sufficient to support the clinicians' use of prescribed MI strategies during intakes. However, the individualized feedback and coaching based on a careful review of audio recorded sessions, as occurred in MIA: STEP, may have specifically helped these clinicians develop more skill in how they delivered MI, particularly advanced one, a finding consistent with recent meta-analyses of MI training studies (de Roten et al., 2013; Schwalbe et al., 2014).

A notable finding was the absence of clinical supervision being provided for intakes at the programs. Only five instances of SAU occurred in this study and only at one site. While clinicians may not have reported some instances of supervision that had occurred, they were prompted monthly by the research staff to indicate if they had received individual or group supervision for any of the intakes they had conducted in the study. Our finding suggests that clinical supervision-as-usual may be virtually no supervision at all, at least for the intake process. A prior study of 340 privately funded US substance abuse treatment centers underscores this possibility in that, across centers, relatively little supervisory time was devoted to promoting the use of several evidence-based practices, including MI (Olmstead, Abraham, Martino, & Roman, 2011). Intakes represent a critical juncture of care for clients in that clinicians must engage clients in a discussion about their substance use, assess related biopsychosocial factors, formulate diagnoses, motivate clients for change, and plan treatment accordingly. Reliance solely on individual clinicians' judgments without supervisory checks and balances for intakes is suboptimal and, moreover, is a missed opportunity for improving clinician practice.

Like our predecessors with similar aims testing other supervision approaches (Watkins, 2011a), we failed to show that MIA: STEP, compared to SAU, significantly improved the percentage of scheduled sessions attended by clients, retention rates, or days of primary substance abstinence. Overall, clients in both conditions stayed in treatment (70% and 50% at 4- and 12-week follow-ups, respectively) and attended scheduled sessions (60% at both time points) at similar rates. Likewise, they increased the number of days in which they remained abstinent from their primary substance (from baseline to follow-up points) by about 50%. Several explanations may account for the congruent findings. First, clients in both conditions received some MI during their intakes, and the gradual differential gain in the quality of MI produced by MIA: STEP over the course of the trial may have been insufficient in magnitude to enhance client retention or outcome. In short, we contrasted MI

with a slightly better version of itself. Second, up to 45 minutes of MI (20–30 minutes up front and 10–15 minutes at the end) were used within intakes that were on average 1.3 hours in length. Thus, the psychosocial assessment of the MI “sandwich” was a substantial component, and it may have diluted differential effects produced by supervision conditions. This very brief MI component of a single intake session, followed by varied treatment content and intensity across programs, likely reduced ability to detect the effect of supervision on retention or outcomes. Supervision of MI sessions, unadulterated by an extended psychosocial assessment or other treatment approaches, or of a multi-session MI intervention might have provided a better venue for studying the effectiveness of MIA: STEP. Third, across conditions baseline levels of client motivation to change their alcohol and drug use were quite high, and about half the clients were court-mandated to treatment. This combination of internal and external motivation to change may have limited the clinical relevance of MI for our study’s client sample. Fourth, supervision researchers have called for getting beyond “the acid test” by putting more emphasis on proximal outcomes that supervision may more directly affect (Frietas, 2002; Holloway & Neufeldt, 1995; Milne, 2014; Wampold & Holloway, 1997; Watkins, 2011). For example, the process of eliciting client language that favors change while resolving client statements that disfavor it has been shown to be a key causal mechanism in MI (Magill et al., 2014). MIA: STEP, relative to SAU, may not have sufficiently enhanced the clinicians’ ability to facilitate this causal mechanism. Since client language was not coded in this study, we cannot rule out this possibility. Studying the degree to which competency-based supervision can improve clinicians’ ability to activate mechanisms directly related to client outcomes within different psychotherapeutic approaches should be the focus of future research.

A final consideration is the cost of delivering MIA: STEP relative to SAU. The incremental cost of MIA: STEP, compared to SAU, as implemented across the 11 sites was \$145,774.00. This excess was mostly driven by the cost of time spent by the MIA: STEP experts in training and consulting with the program supervisors, supervisors reviewing the intakes and preparing feedback, and supervisors and clinicians meeting for supervision, almost none of which happened in SAU. Given that MIA: STEP did not result in better client treatment retention or outcomes than SAU, the feasibility and acceptability of using this intensive and relatively expensive form of supervision as a strategy to implement MI in community treatment programs is questionable. Less costly and more effective innovations for training clinicians and sustaining effective implementation are needed to scale up implementation efforts for MI and other evidence-based treatments. Efforts to develop computer-based natural language processing models for evaluating the integrity in which clinicians use MI (Atkins, Steyvers, Imel, & Smyth, 2014) or interactive simulated virtual world MI training programs (Mitchell, Heyden, Heyden, et al., 2011) are promising moves in this direction. Competency-based supervision approaches such as MIA: STEP might be better suited in clinical training environments (e.g., graduate schools, internships, postdoctoral fellowships) where intensive supervision resources typically are provided.

Several limitations of this study should be noted. For clinicians, retention was at best fair at the post-trial and 16-week follow-up, though consistent with several other MI clinician training trials (e.g., Miller et al., 2004; Moyers et al., 2008). Only 45 minutes of every intake session was rated for MI adherence and competence; clinicians may have used MI consistent

or inconsistent strategies in the unrated middle portion of the sessions, which might have affected the study's clinician outcomes. In addition, the near absence of supervision-as-usual during the trial resulted in unbalanced time spent in supervision between the conditions. Moreover, supervisors self-selecting to learn MIA: STEP potentially biased this condition with more interested and motivated supervisors. For clients, we relied on date of discharge to determine program retention. However, programs varied widely in their policies about how long clients could not attend sessions before being discharged. We also relied on self-report substance use data, rather than urine and breath specimen sample results, which we could not gather more frequently than at the 4- and 12-week assessment points. Confirmation of clients' self-reports with the results of urine and breath specimens lend confidence in the accuracy of the clients' self-reported substance use. Finally, we could not control how randomized clients were assigned to clinicians after their intakes and did not track if they had received any form of treatment from clinicians in the alternative condition or had received MI at the program from other sources. Thus, we cannot rule out these complicating and potentially contaminating effects.

4.1. Conclusions

This study had many strengths, including randomization of both clinicians and clients to conditions, clear operationalization of supervision and significant attention to supervisor MIA: STEP integrity, use of successively supervised cases to allow for examination of the growth in MI skills over time, control for workshop effects, and large and diverse samples of clinicians and clients. In addition, unlike most prior MI training studies, which allowed clinicians to self-select clients for MI demonstrations, this study had clinicians conduct MI with patients as they entered treatment, reducing self-selection bias. Given these strengths, the mixed study findings in the context of the cost of MIA: STEP-guided supervision are not easily dismissed. Implementation strategies such as competency-based supervision requires the same degree of careful study as used to establish treatments as evidence-based and should not be presumed to be effective until so proven.

Acknowledgments

This study was funded by the U.S. National Institute on Drug Abuse (R01 DA1049398 awarded to Steve Martino, with additional support provided by R01 DA034243, P50-DA09241 and U10-DA130038). Its contents are solely the responsibility of the authors and do not necessarily represent the official views of NIDA. The authors acknowledge the invaluable support of program directors (Asher Delorme, Ronald Fleming, John Hamilton, Donna LaPaglia, Lynn Madden, Alan Mathis, Nancy Navarretta, Mary Painter, Joseph Riker, Nicolangelo Scibelli, Joseph Sullivan), study personnel (Wendy Bamatter, Kornelia Balogh, Carol Gordon, Catalina Gorgano, Ashley McCaherty, Brian Pinto, Brandi Wells) and independent raters (Theresa Babuscio, Joanne Corvino, Suzanne Decker, Mary Driscoll, Tami Frankforter, Rebecca Gordon, Christoffer Grant, Lyndia Haff, Leslie Lemert, Jennifer Lovallo, Christine Lozano, Cheryl Seifert, Kasey Spreyer, Kristen Wrocklage, Amanda Zayde) involved in this study.

References

- American Psychological Association. Guidelines for clinical supervision in health service psychology. *American Psychologist*. 2015; 70:33–46. [PubMed: 25581007]
- Atkins DC, Steyvers M, Imel ZE, Smyth P. Scaling up the evaluation of psychotherapy: evaluating motivational interviewing fidelity via statistical text classification. *Implementation Science*. 2014; 9:49. [PubMed: 24758152]

- Baer JS, Ball SA, Campbell BK, Miele GM, Schoener EP, Tracy K. Training and fidelity monitoring of behavioral interventions in multi-site addictions research: A review. *Drug and Alcohol Dependence*. 2007; 87:107–118. [PubMed: 17023123]
- Ball, S.; Martino, S.; Corvino, J.; Morgenstern, J.; Carroll, KM. Independent Tape Rater Guide. 2002. Unpublished psychotherapy tape rating manual
- Ball SA, Martino S, Nich C, Frankforter TL, Van Horn D, Crits-Christoph P, Woody GE, Obert JL, Farentinos C, Carroll KM. Site matters: Motivational enhancement therapy in community drug abuse clinics. *Journal of Consulting and Clinical Psychology*. 2007; 75:556–567. [PubMed: 17663610]
- Bambling M, King R, Raue P, Schweitzer R, Lambert W. Clinical supervision: Its influence on client-rated working alliance and client symptom reduction in the brief treatment of major depression. *Psychotherapy Research*. 2006; 16:317–331.
- Beidas RS, Kendall PC. Training therapists in evidence-based practice: A critical review of studies from a systems-contextual perspective. *Clinical Psychology*. 2010; 17:1–30. [PubMed: 20877441]
- Bernard, JM.; Goodyear, RK. Fundamentals of clinical supervision. 5. Boston, MA: Pearson; 2014.
- Beutler LE, Kendall PC. Introduction to the special section: The case for training in the provision of psychological therapy. *Journal of Consulting and Clinical Psychology*. 1995; 63:179–183. [PubMed: 7751478]
- Borders, LD.; DeKruyf, L.; Fernando, DM.; Glossoff, HL.; Hays, DG.; Page, B.; Welfare, LE. Best practices in clinical supervision. Alexandria, VA: Association for Counselor Education and Supervision; 2011. <http://www.acesonline.net/wpcontent/uploads/2011/10/ACES-Best-Practices-in-clinical-supervision-document-FINAL.pdf> [Accessed October 16, 2015]
- Burke BL, Arkowitz H, Menchola M. The efficacy of motivational interviewing: A meta-analysis of controlled clinical trials. *Journal of Consulting and Clinical Psychology*. 2003; 71:843–861. [PubMed: 14516234]
- Carroll KM, Ball SA, Nich C, Martino S, Frankforter TL, Farentinos C, Kunkel L, Mikulich-Gilbertson S, Morgenstern J, Obert JL, Polcin D, Snead N, Woody GE. Motivational interviewing to improve treatment engagement and outcome in individuals seeking treatment for substance abuse: A multisite effectiveness study. *Drug and Alcohol Dependence*. 2006; 81:301–312. [PubMed: 16169159]
- Carroll KM, Fenton LR, Ball SA, Nich C, Frankforter TL, Shi J, Rounsaville BJ. Efficacy of disulfiram and cognitive behavioral therapy in cocaine-dependent outpatients: A randomized placebo-controlled trial. *Archives of General Psychiatry*. 2004; 61:264–272. [PubMed: 14993114]
- Carroll KM, Martino S, Suarez-Morales L, Ball SA, Miller WR, Añez L, Paris M, Nich C, Frankforter TL, Matthews J, Farentinos C, Szapocznik J. Multisite Randomized Controlled Effectiveness Trial of Motivational Enhancement for Spanish-Speaking Substance Users. *Journal of Consulting and Clinical Psychology*. 2009; 77:993–999. [PubMed: 19803579]
- Carroll KM, Martino S, Rounsaville BJ. No train, no gain? *Clinical Psychology: Science and Practice*. 2010; 17:36–40.
- Carroll KM, Nich C, Rounsaville BJ. Use of observer and therapist ratings to monitor delivery of coping skills treatment for cocaine abusers: Utility of therapist session checklists. *Psychotherapy Research*. 1998; 8:307–320.
- Carroll KM, Nich C, Sifry R, Frankforter T, Nuro KF, Ball SA, Fenton LR, Rounsaville BJ. A general system for evaluating therapist adherence and competence in psychotherapy research in the addictions. *Drug and Alcohol Dependence*. 2000; 57:225–238. [PubMed: 10661673]
- Center for Substance Abuse Treatment. Competencies for substance abuse treatment clinical supervisors. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2007. Technical Assistance Publication (TAP) Services 21-A. DHHS Publication No. (SMA) 07–4243
- Cicchetti D. Guidelines, criteria, and rules of thumb for evaluating normed and standardized instruments in psychology. *Psychological Assessment*. 1994; 6:284–290.
- Condon TP, Miner LL, Balmer CW, Pintello D. Blending addiction research and practice: Strategies for technology transfer. *Journal of Substance Abuse Treatment*. 2008; 35:156–160. [PubMed: 18337054]

- de Roten Y, Zimmermann G, Ortega D, Delpland J. Meta-analysis of the effects of MI training on clinician behavior. *Journal of Substance Abuse Treatment*. 2013; 45:155–62. [PubMed: 23537923]
- Ellis, MV.; Ladany, N. Inferences concerning supervisees and clients in clinical supervision: An integrative review. In: Watkins, CE., Jr, editor. *Handbook of psychotherapy supervision*. New York, NY: Wiley; 1997. p. 447-507.
- Ellis MV, Ladany N, Krenzel M, Schult D. Clinical supervision research from 1981–1993: A methodological critic. *Journal of Counseling Psychology*. 1996; 43:35–50.
- Falender CA, Cornish JAE, Goodyear R, Hatcher R, Kaslow NJ, Leventhal G, et al. Defining competencies in psychological supervision: A consensus statement. *Journal of Clinical Psychology*. 2004; 60:771–785. [PubMed: 15195339]
- Falender CA, Shafranske EP. Competence in competency-based supervision practice: Construct and application. *Professional Psychology: Research and Practice*. 2007; 38:232–240.
- Falender CA, Shafranske EP. The importance of competency-based clinical supervision and training in the twenty-first century: Why bother? *Journal of Contemporary Psychology*. 2012; 42:129–137.
- Fals-Stewart W, O'Farrell TJ, Frietas 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; 68:134–144. [PubMed: 10710848]
- Farber EW, Kaslow NJ. Introduction to the special section: The role of supervision in ensuring the development of psychotherapy competencies across diverse theoretical perspectives. *Psychotherapy: Theory, Research, Practice, Training*. 2010; 47:1–2.
- Feingold A. Effect sizes for growth-modeling analysis for controlled clinical trials in the same metric as for classical analysis. *Psychological Methods*. 2009; 14:43–53. [PubMed: 19271847]
- Fleming, I. Training clinical psychologists as supervisors. In: Fleming, I.; Steen, L., editors. *Supervision and clinical psychology: Theory practice and perspectives*. London: Brunner-Routledge; 2004. p. 72-92.
- Gibbons CJ, Carroll KM, Ball SA, Nich C, Frankforter TL, Martino S. Community program therapist adherence and competence in a motivational interviewing assessment intake session. *American Journal of Drug and Alcohol Abuse*. 2010; 36:342–349. [PubMed: 20942726]
- Gonsalvez CJ, Milne DL. Clinical supervisor training in Australia: A review of current problems and possible solutions. *Australian Psychologist*. 2010; 45:233–242.
- Hettema J, Steele J, Miller WR. Motivational interviewing. *Annual Reviews in Clinical Psychology*. 2005; 1:91–111.
- Herschell AD, Kolko DJ, Baumann BL, Davis AC. The role of therapist training in the implementation of psychosocial treatments: A review and critique with recommendations. *Clinical Psychology Review*. 2010; 30:448–466. [PubMed: 20304542]
- Hoge MA, Migdole S, Farkas MS, Ponce AN, Hunnicutt C. Supervision in public sector public health: A review. *The Clinical Supervisor*. 2011; 30:183–203.
- Holloway EL, Neufeldt SA. Supervision: Its contribution to treatment efficacy. *Journal of Consulting and Clinical Psychology*. 1995; 63:207–213. [PubMed: 7751481]
- Knudsen HK, Ducharme LJ, Roman PM. Clinical supervision, emotional exhaustion, and turnover intention: A study of substance abuse treatment counselors in the Clinical Trials Network of the National Institute on Drug Abuse. *Journal of Substance Abuse Treatment*. 2008; 35:387–395. [PubMed: 18424048]
- Krassler HR, Kadden RM, Babor TF, Tennen H, Rounsaville BJ. Validity of the SCID in substance abuse patients. *Addiction*. 1996; 91:859–868. [PubMed: 8696248]
- Lichtenberg JW, Portnoy SM, Bebeau MJ, Leigh IW, Nelson PD, Rubin NJ, et al. Challenges to the assessment of competence and competencies. *Professional Psychology: Research and Practice*. 2007; 38:474–478.
- Lundahl BW, Kunz C, Brownell C, Tollefson D, Burke B. Meta-analysis of motivational interviewing: Twenty Five years of empirical studies. *Research on Social Work Practice*. 2010; 20:137–160.
- Madson MB, Loignon AC, Lane C. Training in motivational interviewing: a systematic review. *Journal of Substance Abuse Treatment*. 2009; 26:101–109. [PubMed: 18657936]

- Magill M, Gaume J, Apodaca TR, Walthers J, Mastroleo NR, Borsari B, Longabaugh R. The technical hypothesis of motivational interviewing: A meta-analysis of MI's key causal model. *Journal of Consulting and Clinical Psychology*. 82:1–11.
- Mannix KA, Blackburn IM, Garland A, Gracie J, Moorey S, Reid B, et al. Effectiveness of brief training in cognitive behaviour therapy techniques for palliative care practitioners. *Palliative Medicine*. 2006; 20:579–584. [PubMed: 17060250]
- Martino S. Strategies for training counselors in evidence-based treatments. *Addiction Science and Clinical Practice*. 2010; 5:30–40. [PubMed: 22002451]
- Martino, S.; Ball, SA.; Gallon, SL.; Hall, D.; Garcia, M.; Ceperich, S.; Farentinos, C.; Hamilton, J.; Hausotter, W. Motivational interviewing assessment: Supervisory tools for enhancing proficiency. Salem, OR: Northwest Frontier Addiction Technology Transfer Center, Oregon Health and Science University; 2006.
- Martino S, Ball SA, Nich C, Canning-Ball M, Rounsaville BJ, Carroll KM. Teaching community program clinicians motivational interviewing using expert and train-the-trainer strategies. *Addiction*. 2011; 106:428–441. [PubMed: 20925684]
- Martino S, Ball SA, Nich C, Frankforter TL, Carroll KM. Community program therapist adherence and competence in motivational enhancement therapy. *Drug and Alcohol Dependence*. 2008; 96:37–48. [PubMed: 18328638]
- Martino S, Brigham GS, Higgins C, Gallon S, Freese TE, Albright LM, Hulsey EG, Krom L, Storti SA, Perl H, Nugent CD, Pintello D, Condon TP. Partnerships and pathways of dissemination: The NIDA-SAMHSA Blending Initiative in the Clinical Trials Network. *Journal of Substance Abuse Treatment*. 2010; 38:S31–S48. doi. [PubMed: 20307793]
- Martino S, Canning-Ball M, Carroll KM, Rounsaville BJ. A criterion-based stepwise approach for training counselors in motivational interviewing. *Journal of Substance Abuse Treatment*. 2011; 40:357–365. [PubMed: 21316183]
- Martino, S.; Paris, M.; Añez, L.; Carroll, KM. Supervisor adherence and competence scale. 2012. Unpublished supervision audio recording rating manual
- McLellan AT, Alterman AI, Cacciola J, Metzger D, O'Brien CP. A new measure of substance abuse treatment: Initial studies of the Treatment Services Review. *Journal of Nervous and Mental Disease*. 1992; 180:101–110. [PubMed: 1737971]
- Miller, WR.; Rollnick, S. Motivational interviewing: Preparing people for change. 2. New York, NY: Guilford Press; 2002.
- Miller, WR.; Rollnick, S. Motivational interviewing: Helping people change. 3. New York, NY: Guilford Press; 2013.
- Miller WR, Rose GS. Toward a theory of motivational interviewing. *American Psychologist*. 2009; 64:527–537. [PubMed: 19739882]
- Miller WR, Sorensen JL, Selzer JA, Brigham GS. Disseminating evidence-based practices in substance abuse treatment: a review with suggestions. *J Subst Abuse Treat*. 2006; 31(1):25–39. [PubMed: 16814008]
- Miller WR, Yahne CE, Moyers TB, Martinez J, Pirritano M. A randomized trial of methods to help clinicians learn motivational interviewing. *Journal of Consulting and Clinical Psychology*. 2004; 72:1050–1062. [PubMed: 15612851]
- Milne DL. Beyond the “Acid Test”: a conceptual review and reformulation of outcome evaluation in clinical supervision. *American Journal of Psychotherapy*. 2014; 68:213–230. [PubMed: 25122986]
- Milne DL, James IA. A systematic review of effective cognitive-behavioural supervision. *British Journal of Clinical Psychology*. 2000; 39:111–127. [PubMed: 10895356]
- Milne DL, Sheikh AI, Pattison S, Wilkinson A. Evidence-based training for clinical supervisors: A systematic review of 11 controlled studies. *The Clinical Supervisor*. 2011; 30:53–71.
- Miller WR, Tonigan JS. Assessing drinkers' motivation for change: The Stages of Change Readiness and Treatment Eagerness Scale (SOCRATES). *Psychology of Addictive Behaviors*. 1996; 10:81–89.
- Mitchell S, Heyden R, Heyden N, Schroy P, Andrew S, Sadikova E, Wiecha J. A pilot study of motivational interviewing training in a virtual world. *Journal of Medical Internet Research*. 2011; 13:277–286.

- Olmstead TA, Abraham AJ, Martino S, Roman PM. Counselor training in several evidence-based psychosocial addiction treatments in private US substance abuse treatment centers. *Drug and Alcohol Dependence*. 2012; 120:149–154. [PubMed: 21831536]
- Olmstead T, Canning-Ball M, Carroll KM, Martino S. Cost and cost-effectiveness of three strategies for training clinicians in motivational interviewing. *Drug and Alcohol Dependence*. 2011; 116:195–202. [PubMed: 21277713]
- Patton M, Kivlighan D. Relevance of the supervisory alliance to the counseling alliance and to treatment adherence in counselor training. *Journal of Counseling Psychology*. 1997; 44:149–154.
- Raudenbush SW. Statistical analysis and optimal design for cluster randomized trials. *Psychological Methods*. 1997; 2:173–185.
- Raudenbush SW, Liu X. Effects of study duration, frequency of observation, and sample size on power in studies of group differences in polynomial change. *Psychological Methods*. 2001; 6:387–401. [PubMed: 11778679]
- Raudenbush SW, Xiaofeng L. Statistical power and optimal design for multisite randomized trials. *Psychological Methods*. 2003; 5:199–213. [PubMed: 10937329]
- Rakovshik S, McManus G. Establishing evidence-based training in cognitive behavioral therapy: A review of current empirical findings and theoretical guidance. *Clinical Psychology Review*. 2010; 30:496–516. [PubMed: 20488599]
- Reisner RP, Milne D. Supervising cognitive-behavioral psychotherapy: Pressing needs, impressing possibilities. *Journal of Contemporary Psychotherapy*. 2012; 42:161–171.
- Reisner RP, Milne D. A systematic review and reformulation of outcome evaluation in clinical supervision: Applying the fidelity framework. *Training and Education in Professional Psychology*. 2014; 8:149–157.
- Roche AM, Todd CL, O'Connor J. Clinical supervision in the alcohol and other drugs field: An imperative or an option? *Drug and Alcohol Review*. 2007; 26:241–249. [PubMed: 17454013]
- Roth AD, Pilling S, Turner J. Therapist training and supervision in clinical trials: Implications for clinical practice. *Behavioural and Cognitive Psychotherapy*. 2010; 38:291–302. [PubMed: 20367895]
- Santa Ana EJ, Carroll KM, Anez L, Paris M, Ball SA, Nich C, Frankforter TL, Suarez-Morales L, Szapocznik J, Martino S. Evaluating motivational enhancement therapy adherence and competence among Spanish-speaking therapists. *Drug and Alcohol Dependence*. 2009; 103:44–51. [PubMed: 19394164]
- Schoenwald SK, Mehta TG, Frazier SL, Shernoff ES. Clinical supervision in effectiveness and implementation research. *Clinical Psychology Science and Practice*. 2013; 20:44–59.
- Schoenwald SK, Sheidow AJ, Chapman JE. Clinical supervision in treatment transport: Effects on adherence and outcomes. *Journal of Consulting and Clinical Psychology*. 2009; 77:410–421. [PubMed: 19485583]
- Sholomskas DE, Syracuse-Siewert G, Rounsaville BJ, Ball SA, Nuro KF, Carroll KM. We don't train in vain: A dissemination trial of three strategies of training clinicians in cognitive behavioral therapy. *Journal of Consulting and Clinical Psychology*. 2005; 73:106–115. [PubMed: 15709837]
- Schwalbe CS, Hans YO, Zweben A. Sustaining motivational interviewing: A meta-analysis of training studies. *Addiction*. 2009; 104:1287–1294. [PubMed: 24661345]
- Smedslund G, Berg RC, Hammerstrom KT, Steiro A, Leiknes KA, Dahl HM, Karlsen K. Motivational interviewing for substance abuse. *Cochrane Database of Systematic Reviews*. 2011; (5) Art. No.: CD008603.
- Smith JS, Carpenter KM, Amrhein P, Brooks A, Levin D, Schreiber EA, et al. Training substance abuse clinicians in motivational interviewing using live supervision via tele-conferencing technology. *Journal of Consulting and Clinical Psychology*. 2012; 80:450–464. [PubMed: 22506795]
- Sobell LC, Sobell MB. Timeline followback: A technique for assessing self-reported alcohol consumption. In: Litten, RZ.; Allen, J., editors. *Measuring alcohol consumption: Psychosocial and biological methods*. New Jersey: Humana Press; 1992. p. 41–72.

- Soderlund LL, Madson MB, 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]
- Stout RL, Wirtz PW, Carbonari JP, Delboca FK. Ensuring balanced distribution of prognostic factors in treatment outcome research. *Journal on Studies of Alcoholism*. 1994; 12(Suppl):70–75.
- Walters ST, Matson SA, Baer JS, Zeidonis DM. Effectiveness of workshop training for psychosocial addiction treatments: A systematic review. *Journal of Substance Abuse Treatment*. 2005; 29:283–293. [PubMed: 16311181]
- Watkins CE. Does psychotherapy supervision contribute to patient outcomes? Considering thirty years of research. *The Clinical Supervisor*. 2011a; 30:235–256.
- Watkins CE. Psychotherapy supervision since 1909: Some friendly observations about its first century. *Journal of Contemporary Psychotherapy*. 2011b; 41:57–67.
- Watkins CE, Scaturro DJ. Toward an integrative, learning-based model of psychotherapy supervision: Supervisory alliance, educational interventions, and supervisee learning/relearning. *Journal of Psychotherapy Integration*. 2013; 23:75–95.
- Wheeler S, Richards K. The impact of clinical supervision on counsellors and therapists, their practice and their clients: A systematic review of the literature. *Counselling and Psychotherapy Research*. 2007; 7:54–65.

Highlights

- The effectiveness of MIA: STEP supervision for motivational interviewing was determined.
- Clinicians supervised with MIA: STEP significantly increased their MI competency.
- MIA: STEP supervision did not improve client treatment retention or outcomes.
- MIA: STEP was more expensive to deliver than usual supervision practices.
- Innovative alternatives to resource-intensive supervision approaches are needed.

Figure 1. CONSORT diagram of clinician eligibility, enrollment, randomization, intakes completed, and follow-up rates. MIA: STEP = Motivational Interviewing Assessment: Supervisory Tools for Enhancing Proficiency; SAU = supervision-as-usual.

Figure 2. CONSORT diagram of client eligibility, enrollment, randomization, completed intake, and follow-up rates. MIA: STEP = Motivational Interviewing Assessment: Supervisory Tools for Enhancing Proficiency; SAU = supervision-as-usual.

Table 1

Clinician baseline demographic variables and experience levels by supervision condition.

Variable	Condition	
	MIA: STEP n = 35	SAU n = 31
Female, <i>n</i> (%)	30 (86%)	22 (71%)
Mean (SD) age	39.2 (12.7)	43.6 (13.6)
Ethnicity, <i>n</i> (%)		
African-American	4 (11%)	5 (16%)
Caucasian-American	21 (60%)	22 (71%)
Hispanic-American	9 (26%)	4 (13%)
Multiracial	1 (3%)	0
Mean (SD) Years Education	17.7 (2.8)	17.7 (2.1)
Highest Degree Earned, <i>n</i> (%)		
Doctorate	1 (3%)	0
Master's	24 (69%)	21 (68%)
Bachelor's	4 (11%)	5 (16%)
Associates	2 (6%)	3 (10%)
High School	4 (11%)	2 (6%)
Licensure Status, <i>n</i> (%)		
Alcohol/Drug Certification	13 (37%)	16 (52%)
Licensed Professional	17 (49%)	17 (55%)
Mean (SD) Years Experience		
General Counseling	7.5 (7.0)	10.4 (10.1)
Substance Abuse Counseling	6.1 (6.9)	9.4 (10.3)
Past year MI training		
Mean (SD) hours	2.3 (7.2)	2.6 (4.9)
Lecture/presentation, <i>n</i> (%)	7 (20%)	1 (3%)
Workshop, <i>n</i> (%)	3 (9%)	1 (3%)
Supervision, <i>n</i> (%)	1 (3%)	0 (0%)
Read MI Textbook	30 (86%)	21 (68%)
Used MI Manuals	9 (26%)	9 (29%)

Note. MIA: STEP = Motivational Interviewing Assessment: Supervisory Tools for Enhancing Proficiency, SAU = Supervision-as-Usual. For the Licensure Status category, the percentages in the SAU condition exceed 100% because one clinician was certified in alcohol/drug counseling and licensed in her discipline. There are no significant condition differences on the variables.

Table 2

Client baseline demographic and substance use variables by supervision condition.

Variable	Condition	
	MIA: STEP n = 227	SAU n = 223
Female, <i>n</i> (%)	75 (33%)	78 (35%)
Mean (SD) age	35.5 (11.8)	35.4 (11.5)
Ethnicity, <i>n</i> (%)		
African-American	111 (49%)	96 (43%)
Caucasian-American	59 (26%)	66 (30%)
Hispanic-American	47 (21%)	48 (22%)
Multiracial	6 (3%)	9 (4%)
Other	1 (1%)	4 (2%)
Mean (SD) Years Education	12.0 (2.4)	12.2 (2.4)
Employment Status, <i>n</i> (%)		
Full-time	29 (13%)	30 (17%)
Part-time	38 (17%)	42 (19%)
Unemployed	160 (70%)	143 (64%)
Marital Status, <i>n</i> (%)		
Single	141 (62%)	133 (60%)
Married or cohabitating	41 (18%)	40 (18%)
Divorced, separated or widowed	45 (20%)	50 (22%)
Primary Substance Used Past 4 weeks, <i>n</i> (%)		
Alcohol	69 (30%)	73 (33%)
Cannabis	80 (35%)	75 (34%)
Cocaine	55 (24%)	51 (23%)
Opioids	13 (6%)	16 (7%)
Stimulants	1 (0.5%)	2 (1%)
Benzodiazepines	1 (0.5%)	0
Other	8 (4%)	6 (3%)
Mean (SD) Days Primary Substance Abstinence	14.8 (9.8)	14.6 (10.2)
Any Substance Used Past 4 Weeks, <i>n</i> (%)		
Alcohol	188 (83%)	168 (75%)
Cannabis	129 (57%)	128 (57%)
Cocaine	85 (37%)	84 (38%)
Opioids	34 (15%)	35 (16%)
Hallucinogens/PCP	25 (11%)	23 (10%)
Stimulants	4 (2%)	2 (1%)
Inhalants	1 (0.5%)	1 (0.5%)
Current DSM-IV Substance Use Disorder, <i>n</i> (%)		
Alcohol	89 (39%)	88 (39%)
Cannabis	78 (34%)	79 (35%)
Cocaine	62 (27%)	62 (28%)

Variable	Condition	
	MIA: STEP n = 227	SAU n = 223
Opioids	14 (6%)	18 (8%)
Hallucinogens/PCP	10 (4%)	5 (2%)
Stimulants	1 (0.5%)	2 (1%)
Court-Mandated to Treatment	110 (48%)	112 (50%)
Mean Socrates – Primary Alcohol		
Recognition	26.9 (6.6)	26.6 (8.0)
Ambivalence	15.1 (3.7)	14.2 (4.4)
Taking Steps	31.9 (5.4)	31.1 (6.9)
Mean Socrates – Primary Drug		
Recognition	26.8 (7.9)	27.2 (7.3)
Ambivalence	14.2 (4.4)	13.8 (4.4)
Taking Steps	31.5 (6.4)	32.2 (6.2)

Note. MIA: STEP = Motivational Interviewing Assessment: Supervisory Tools for Enhancing Proficiency, SAU = Supervision-as-Usual. For the Any Substance Used Past 4 Weeks and DSM-IV Substance Use Disorder categories, the percentages in the both conditions exceed 100% because one clients could have been using multiple substances and have several concurrent disorders. There are no significant condition differences on the variables.

Table 3

MIA: STEP and SAU supervision adherence and competence by supervision condition and site.

Scale	Supervision Condition			Site	Condition x Site
	MIA: STEP	SAU	Supervision Condition		
MIA: STEP Consistent Strategies					
Adherence					
<i>M</i>	4.67	2.27	<i>F</i> (1, 209)=108.52 <i>p</i> =.000	<i>F</i> (10, 209) =3.36 <i>p</i> =.000	<i>F</i> (10, 209) =1.16 <i>p</i> =.320
<i>SD</i>	1.60	1.22			
<i>N</i>	155	55			
Competence					
<i>M</i>	5.05	3.72	<i>F</i> (1, 190) =42.88 <i>p</i> =.000	<i>F</i> (10, 190) =4.86 <i>p</i> =.000	<i>F</i> (10, 190) =.757 <i>p</i> =.670
<i>SD</i>	1.19	1.40			
<i>N</i>	148	43			
General Supervision Strategies					
Adherence					
<i>M</i>	1.91	2.88	<i>F</i> (1, 209) =62.64 <i>p</i> =.000	<i>F</i> (10, 209) =2.30 <i>p</i> =.014	<i>F</i> (10, 209) =1.21 <i>p</i> =.286
<i>SD</i>	.79	.67			
<i>N</i>	155	55			
Competence					
<i>M</i>	3.96	3.91	<i>F</i> (1, 174) =.026 <i>p</i> =.873	<i>F</i> (10, 174) =2.25 <i>p</i> =.018	<i>F</i> (10, 174) =1.14 <i>p</i> =.334
<i>SD</i>	1.11	1.00			
<i>N</i>	120	55			
MI Supervision Style					
<i>M</i>	4.12	2.00	<i>F</i> (1, 209) =64.24 <i>p</i> =.000	<i>F</i> (10, 209) =3.06 <i>p</i> =.001	<i>F</i> (10, 209) =1.41 <i>p</i> =.178
<i>SD</i>	1.84	1.55			
<i>N</i>	155	55			

Note. MIA: STEP = Motivational Interviewing Assessment: Supervisory Tools for Enhancing Proficiency, SAU = Supervision-as-Usual. Adherence was rated 1 = not at all, 2 = a little, 3 = infrequently, 4 = somewhat, 5 = quite a bit, 6 = considerably, 7 = extensively. Competence was rated 1 = very poor, 2 = poor, 3 = acceptable, 4 = adequate, 5 = good, 6 = very good, 7 = excellent. A familywise alpha of $p < .01$ (.05/5) marked significant effects.

Table 4

Clinician MI adherence and competence by supervision condition and trial case.

Variable	Supervision Condition			Supervision Condition	Trial Case	Condition by Trial Interaction
	MIA: STEP M (SD)	SAU M (SD)				
Post-trial outcomes						
Fundamental MI Adherence						
Baseline	3.71 (.87)	3.55 (.78)	F	8.16	8.22	3.12
Post-Workshop	4.43 (.70)	4.07 (.72)	p	.01	.01	.08
1	4.84 (.59)	4.27 (.96)	df	1, 114	1, 238	1, 468
2	4.68 (.93)	4.16 (.74)	d	.53	.08	.54
3	4.53 (.81)	4.08 (.88)				
4	4.94 (.79)	4.06 (.86)				
5	4.65 (.84)	4.08 (.77)				
6	4.46 (.97)	3.90 (1.03)				
7	4.33 (.61)	3.87 (.74)				
16-week follow-up	4.73 (.59)	3.66 (1.18)	F	7.21	8.22	6.44
			p	.01	.01	.11
			df	1, 113	1, 228	1, 512
			d	.33	.05	.46
Fundamental MI Competence						
Baseline	3.57 (.75)	3.74 (.69)	F	1.76	9.29	5.73
Post-Workshop	4.14 (.68)	3.91 (.79)	p	.19	.00	.01
1	4.36 (.72)	4.04 (.74)	df	1, 137	1, 480	1, 521
2	4.21 (.83)	3.84 (.66)	d	.20	.08	.63
3	4.28 (.84)	3.98 (.68)				
4	4.42 (.89)	3.90 (.58)				
5	4.15 (.71)	3.79 (.57)				
6	4.18 (.65)	3.90 (.80)				
7	4.35 (.74)	3.82 (.61)				
16-week follow-up	4.47 (.65)	3.87 (.63)	F	1.87	11.63	7.18
			p	.17	.00	.01

Variable	Supervision Condition		Supervision Condition	Trial Case	Condition by Trial Interaction		
	MIA: STEP M (SD)	SAU M (SD)					
Advanced MI Adherence			df	1, 169	1, 524	1, 521	
			d	.20	.08	.72	
	Baseline		F	4.51	6.95	2.46	
	Post-Workshop		p	.04	.01	.12	
	1	2.01 (.78)	2.03 (.83)	df	1, 120	1, 125	1, 469
	2	2.81 (.79)	2.54 (.85)	d	.33	.10	.43
	3	3.56 (1.04)	2.97 (1.25)				
	4	3.25 (.97)	3.03 (1.20)				
	5	3.51 (1.26)	2.43 (.95)				
	6	3.25 (1.11)	2.77 (1.21)				
	7	3.47 (1.27)	2.72 (.96)				
		3.17 (1.10)	2.68 (1.18)				
		3.01 (.88)	2.65 (1.07)				
	16-week follow-up			F	4.23	9.07	4.45
	3.23 (.75)	2.19 (.65)	p	.04	.00	.04	
			df	1,122	1,238	1,516	
			d	.36	.10	.65	
Advanced MI Competence			F	0.004	16.09	12.23	
	Baseline		p	.95	.00	.00	
	Post-Workshop		df	1, 126	1, 272	1, 459	
	1	4.24 (.64)	4.09 (.66)	d	.01	.09	.84
	2	4.28 (.69)	4.12 (.57)				
	3	4.24 (.73)	3.98 (.54)				
	4	4.23 (.57)	3.88 (.45)				
	5	4.03 (.65)	3.80 (.66)				
	6	4.14 (.45)	4.00 (.45)				
	7	4.35 (.64)	3.71 (.54)				
	16-week follow-up			F	.02	16.14	11.31
		4.35 (.58)	3.96 (.71)	p	.89	.00	.00
				df	1, 141	1, 511	1, 507
				d	.16	.08	.82

Variable	Supervision Condition		Supervision Condition	Trial Case	Condition by Trial Interaction
	MIA: STEP <i>M (SD)</i>	SAU <i>M (SD)</i>			
MI Inconsistent Adherence					
Baseline	1.53 (.43)	1.66 (.56)	<i>F</i>	.13	.04
Post-Workshop	1.39 (.40)	1.79 (.78)	<i>p</i>	.72	.84
1	1.53 (.45)	1.71 (.53)	<i>df</i>	1, 79	1, 76
2	1.55 (.54)	1.83 (.56)	<i>d</i>	.34	.05
3	1.75 (.61)	1.69 (.54)			
4	1.54 (.51)	1.84 (.55)			
5	1.55 (.53)	1.82 (.81)			
6	1.40 (.37)	1.57 (.39)			
7	1.62 (.47)	1.63 (.65)			
16-week follow-up			<i>F</i>	4.86	.00
			<i>p</i>	.03	.97
			<i>df</i>	1, 80	1, 71
			<i>d</i>	.33	.00

Note. MIA: STEP = Motivational Interviewing Assessment: Supervisory Tools for Enhancing Proficiency, SAU = Supervision-as-Usual. Adherence was rated 1 = not at all, 2 = a little, 3 = infrequently, 4 = somewhat, 5 = quite a bit, 6 = considerably, 7 = extensively. Competence was rated 1 = very poor, 2 = poor, 3 = acceptable, 4 = adequate, 5 = good, 6 = very good, 7 = excellent. The random effect of site is nested in training condition within the hierarchical linear regression models. A familywise alpha of $p < .01$ (.05/5) marked significant effects.

Table 5

Client program retention and primary substance use outcomes by supervision condition and site.

	Supervision Condition			
Variable	MIA: STEP	SAU	Supervision Condition	Site
4-week outcomes				
Percent sessions attended				
<i>M</i>	60.82	59.27	<i>F</i> (1, 374) = .21 <i>p</i> = .65	<i>F</i> (1, 391) = 1.30 <i>p</i> = .20
<i>SD</i>	32.36	32.55	<i>partial eta</i> ² = .004	<i>partial eta</i> ² = .05
Percent retained in treatment	73.0%	70.7%	<i>Wald</i> (1) = .29 <i>p</i> = .59	<i>Wald</i> (1) = .19 <i>p</i> = .66
Days of primary substance abstinence				
<i>M</i>	21.96	22.23	<i>F</i> (1, 370) = .01 <i>p</i> = .94	<i>F</i> (15, 370) = 2.56 <i>p</i> = .001
<i>SD</i>	8.28	8.33	<i>partial eta</i> ² = .000	<i>partial eta</i> ² = .09
12-week outcomes				
Percent sessions attended				
<i>M</i>	60.83	57.85	<i>F</i> (1, 284) = .82 <i>p</i> = .37	<i>F</i> (15, 284) = 1.13 <i>p</i> = .33
<i>SD</i>	41.48	29.01	<i>partial eta</i> ² = .01	<i>partial eta</i> ² = .06
Percent retained in treatment	51.5. %	52.9%	<i>Wald</i> (1) = .14 <i>p</i> = .71	<i>Wald</i> (1) = 4.71 <i>p</i> = .03
Days of primary substance abstinence				
<i>M</i>	22.35	22.80	<i>F</i> (1, 353) = 0.01 <i>p</i> = .94	<i>F</i> (15, 353) = 3.26 <i>p</i> = .00
<i>SD</i>	8.34	7.82	<i>partial eta</i> ² = .00	<i>partial eta</i> ² = .12

Note. MIA: STEP = Motivational Interviewing Assessment: Supervisory Tools for Enhancing Proficiency, SAU = Supervision-as-Usual. A family-wise alpha of $p < .017$ (.05/3) marked significant effects.

Table 6

Overall cost (in 2014 dollars) by supervision condition.

Cost Category	Supervision Condition	
	MIA: STEP (\$)	SAU (\$)
MI Workshops ^a	34,958	18,236
MIA: STEP Workshop ^b	21,592	0
Post-Workshop Practice Case ^c	8,276	0
Consultations ^d	56,077	0
Supervisions ^e	40,676	151
Self-study MI Materials ^f	12,261	9680
Miscellaneous ^g	758	758
Total	174,599	28,825

Notes. MIA: STEP = Motivational Interviewing Assessment: Supervisory Tools for Enhancing Proficiency, SAU = Supervision-as-Usual. Incremental cost of MIA: STEP, compared to SAU: \$145,774.00.

^aThere were 18 MI workshops, including seven extra MI workshops that were required as 11 clinicians (2 SAU, 9 MIA: STEP) were replaced during the trial.

^bThere were 4 MIA: STEP workshops.

^cThere were 11 post-workshop practice cases (one for each supervisor assigned to MIA: STEP).

^dThere were 97 consultations between experts and supervisors.

^eThere were 5 supervisions in SAU and 247 supervisions in MIA: STEP.

^fIncludes self-study time during company time only.

^gMiscellaneous costs (e.g., DVD player, LCD projector) split equally between supervision conditions.