

Sustaining motivational interviewing: a meta-analysis of training studies

Craig S. Schwalbe, Hans Y. Oh & Allen Zweben

Columbia University School of Social Work, New York, NY, USA

ABSTRACT

Background and Aims Previous research indicates that motivational interviewing (MI) skills decline over time among participants in training workshops when post-workshop feedback and coaching are not provided. This study explored moderators of skill retention among trainees learning MI mainly for substance use disorder treatment in real-world treatment settings, including workshop enhancements and type and dose of post-workshop feedback and coaching. **Methods** A meta-analysis of training studies was conducted with studies that reported MI skills using observational measures and that included trainees from real-world agency settings. Standardized change scores were calculated to indicate the magnitude of pre–post training change in MI skills; standardized change scores from post-training to 3 and 6+ months follow-up were calculated to indicate the sustainability of training gains over time. Effect sizes were aggregated using random effects models. **Results** Twenty-one papers that reported the effects of MI training on agency staff were included in this review. Across studies, training yielded gains in MI skills ($d = 0.76$). Studies that did not include feedback and/or coaching reported eroding skills over a 6-month follow-up ($d = -0.30$), whereas post-workshop feedback/coaching sustained skills ($d = 0.03$). Effects of post-workshop feedback/coaching were moderated by frequency, duration and length of training. Moreover, studies reporting low levels of attrition from training protocols showed small increases in skills over the 6-month follow-up period ($d = 0.12$), whereas studies with high attrition showed skill erosion ($d = -0.29$). **Conclusions** On average, three to four feedback/coaching sessions over a 6-month period sustain skills among trainees for motivational interviewing, mainly for substance use disorder treatment. However, high rates of attrition from feedback/coaching contributes to post-workshop skill erosion.

Keywords Feedback/coaching, meta-analysis, motivational interviewing, supervision, sustainability, training.

Correspondence to: Craig S. Schwalbe, Columbia University School of Social Work, 1255 Amsterdam Avenue, New York, NY 10027, USA.
E-mail: css2109@columbia.edu

Submitted 19 September 2013; initial review completed 22 November 2013; final version accepted 18 March 2014

INTRODUCTION

As human services practitioners and organizations adopt and implement motivational interviewing (MI), the field stands in need of training strategies to foster its long-term sustainability in as-usual treatment settings. Most MI scholars and trainers recognize that reading the MI literature (e.g. manuals) and participating in training workshops are not sufficient to sustain training gains for most human service professionals. Rather, evidence from practitioner training in evidence-based behavioral interventions [1], along with experimental studies from the MI training literature [2], support a multi-modal training approach. However, there has not been a systematic effort to date to establish the minimal training requirements needed to sustain MI skills over extended periods of time.

The current study addressed this gap through a meta-analysis of the MI training literature.

Findings of prior systematic reviews

MI training studies have been the subject of two prior systematic reviews and one meta-analysis [3–5]. The two systematic reviews identify trends in MI training and research [3,4]. Both noted that MI skill increases are observed in most studies, and both concluded that post-workshop inputs such as feedback and coaching are needed to sustain training gains over time. The meta-analysis found that MI training has an overall positive effect on clinician skills at post-test [effect size (ES) = 0.70] and follow-up (ES = 0.69) [5]. The utilization of post-workshop coaching and feedback was

associated with slightly higher effect sizes (ES = 0.82 versus ES = 0.69). None of these reviews attempted to isolate training features, such as the structure of training workshops or the dose and structure of post-workshop training inputs, which may be associated with maintenance of MI skills over time.

Can enhancements to the standard workshop model increase the sustainability of training effects?

The conventional wisdom of the field is that MI skills erode quickly when skills-building workshops are not followed by coaching and feedback. Several studies have demonstrated this empirically, showing diminished skills among workshop participants at follow-up times as short as 2 months [2,6,7]. Whether enhancements or structural changes to a typical training workshop can influence the sustainability of MI skills has been a seldom-asked question. Usually, training workshops include a mix of didactic presentation, demonstrations and practice delivered over 1–3 days [4]. To sustain skills over time, a training workshop needs to influence mediating processes that provide ongoing support to counselor skillfulness, such as organizational support and counselor acceptance [8–10]. Whether workshop enhancements such as Baer's context-tailored training [11] can foster sustainability has not been directly examined.

What dose and type of post-workshop training inputs is required to sustain MI skills over time?

The current literature has little to say about how post-workshop training inputs sustain skills over time. Because few studies have experimentally manipulated feedback and coaching models directly, little can be said about the type and intensity of post-workshop training inputs needed to optimize training outcomes over time. Two notable exceptions are Miller *et al.*'s landmark Evaluating Methods for Motivational Enhancement Education (EMMEE) training study [2] and Smith *et al.*'s evaluation of teleconferencing supervision compared to traditional feedback/coaching [6], both of which found no significant differences in MI skill maintenance between types of post-workshop training inputs deployed. Most other studies have combined individualized feedback from recorded practice samples with coaching or supervision, and to our knowledge no studies have tested the effects of coaching/feedback frequency or duration on long-term outcomes.

Current study

This meta-analysis aimed to fill gaps in the literature about training and support features associated with the maintenance of MI skills over time. Counselors, supervi-

sors and agency administrators need data about effective training programs to inform their decisions about the type of training programs that will foster the acquisition and maintenance of MI skills. Three research questions guided the study:

- 1 Do enhancements to the traditional workshop model of MI training contribute to the maintenance of MI skills over time?
- 2 What types of post-workshop training inputs are associated with MI skill maintenance over time?
- 3 What dose of post-workshop training inputs is needed to sustain MI skills over time?

METHODS

Sample

A search was conducted in three databases (Psycinfo, Medline, ISI Web of Science). The time-period of the search extended to articles published in August 2013. Search terms included 'motivational interviewing' and combinations of the following terms (and their synonyms): implementation, training, translation, fidelity, adaptation, education, sustainability, supervision. We also examined the website of the Motivational Interviewing Network of Trainers (<http://www.motivationalinterview.org/>) and the reference lists of systematic reviews [3–5]. Inclusion criteria included: (i) trainees were human service and medical service professionals from as-usual treatment settings; (ii) outcome measures included direct observational measures of MI skills; and (iii) sufficient information reported to calculate change scores. Studies of student trainees were excluded.

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagram (Fig. 1) [12] shows that 1735 article titles and abstracts were examined and that 62 were selected for more detailed evaluation. Twenty-one studies were included in the final sample (see Table 1).

Coding

The first and second authors coded the studies using a collaborative process. Codes were reviewed weekly and discrepancies were negotiated by consensus. Study authors were contacted as needed for information to calculate effect sizes. Three types of data were coded from the studies, as follows.

Training components

Categorical descriptions of workshop types and post-workshop training inputs were derived inductively from the data. Workshop types included standard workshops, and workshop enhancements that included components

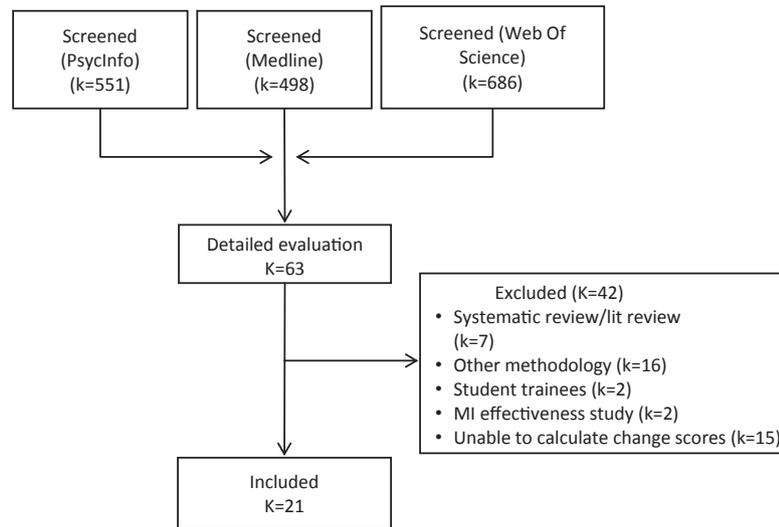


Figure 1 Flow diagram of the study selection process

Table 1 Characteristics of included studies.

Reference	Groups	Design	Training	n	Host site	Setting	%MA	Outcome	Follow months
Baer <i>et al.</i> 2004 [13]	1	Case	Wksp(s)	22	Agency	SUD	78	MISC	2
Baer <i>et al.</i> 2009 [11]	2	RCT	Wksp(s), Wksp(i)	123	Agency	SUD	30	MITI	3
Bennett <i>et al.</i> 2007 [14]	2	RCT	Wksp(i), f/c	54	Center	SUD	–	MITI	4
Bohman <i>et al.</i> 2013 [15]	1	Case	Wksp(s),f/c	36	Agency	Health	0	MITI	8
Britt <i>et al.</i> 2010 [16]	1	Case	Wksp(s), f/c	2	Agency	Health	0	MISC	Post
Brug <i>et al.</i> 2007 [17]	2	RCT	None, Wksp(i), c	37	Agency	Health	0	MISC	Post, 6
Evangelini <i>et al.</i> 2009 [18]	1	Case	Wksp(i),	14	Agency	HIV	0	MITI	Post
Evangelini <i>et al.</i> 2011 [19]	1	Case	Wksp(i)	10	Agency	HIV	0	MITI	12
Forsberg <i>et al.</i> 2010 [20]	1	Case	Wksp(s), f/c	3	Agency	SUD	–	MITI	Post–30
Hartzler <i>et al.</i> 2007 [21]	1	Case	Wksp(i)	23	Agency	SUD	17	MITI	Post
Martino <i>et al.</i> 2011 [22]	3	RCT	SS, Wksp(s), Wksp(i), f/c	92	Agency	SUD	65	ITRS	Post, 3, 6
Martino <i>et al.</i> 2011 [23]	1	RCT	SS, Wksp(i), f/c	26	Agency	SUD	44	ITRS	2, 4, 6
Mastroleo <i>et al.</i> 2010 [24]	2	RCT	Wksp(s), f/c	19	Agency	SUD	0	PEPA	Post
Miller <i>et al.</i> 2001 [25]	1	Case	Wksp(s)	22	Agency	Corr	68	MISC	Post, 4
Miller <i>et al.</i> 2004 [2]	5	RCT	None, SS, Wksp(i), f, c, f/c	140	Center	SUD	85	MISC	Post, 4, 8, 12
Moyers <i>et al.</i> 2008 [7]	3	RCT	None, Wksp(s), f/c	129	Agency	SUD	56	MITI	4, 8, 12
Shafer <i>et al.</i> 2004 [26]	1	Case	Wksp(i)	9	Center	SUD	60	MISC	4
Smith <i>et al.</i> 2007 [27]	1	Case	Wksp(s), f	13	Agency	SUD	38	MITI	Post, 3
Smith <i>et al.</i> 2012 [6]	3	RCT	Wksp(s), f/c	77	Agency	SUD	40	MITI	Post, 2, 5
Van Eijk-Hustings <i>et al.</i> 2010 [28]	1	Case	Wksp(i), f/c	10	Agency	Health	–	MITI	6
Walters <i>et al.</i> 2010 [29]	3	RCT	None, Wksp(i), c	23	Agency	Corr	–	MITI	2, 6

RCT = random clinical trial; Case = case–control design; none = no-training; SS = self-study; Wksp(s) = standard workshop; Wksp(i) = workshop with at least one innovative element (e.g. feedback on training tapes, context-tailored); f = feedback; c = coaching; f/c = feedback and coaching combined; host site = the location of training (Agency = agency-based training, Center = training center); setting = the primary treatment focus of participating practitioners (SUD = substance abuse treatment; Health = health-care clinic setting; Corr = correctional agency); %MA = percentage of trainee group with masters' degree-level education or higher; MISC = Motivational Interviewing Skills Code [30]; MITI = Motivational Interviewing Treatment Integrity Code [31]; ITRS = Independent Tape Rater Scale [22]; PEPA = Peer Proficiency Assessment [24]; follow months = the length of follow-up measures in months.

such as feedback on audio-recorded client sessions, booster sessions (group follow-up sessions separate from feedback/coaching), teleconferencing or context-tailored training courses. Post-workshop inputs included feedback on audio-recorded client interactions and coaching. The duration of training, number of post-training input hours and number of training contacts

were also coded. Post-workshop inputs were estimated to be 1 hour per contact when duration was unspecified.

Methodological moderators

Methodological characteristics included sample size, length of follow-up (post-training, 3 months post-training, 6 months post-training, more than 6 months),

observational outcome measure, outcome reliability intraclass correlation (ICC), study type (experimental with or without random assignment, case-control) and whether a simulated patient or real client interaction was recorded and coded. Setting characteristics [e.g. substance use disorder (SUD) treatment program, criminal justice program] and trainee characteristics (e.g. level of education) were also coded.

Effect sizes

An analysis of change scores treating each training group as a within-subjects design was selected for this study. Although meta-analyses of between-groups designs are frequently reported and preferred, an analysis of within-subjects designs can be justified for conceptual reasons based on the strength of the within-subjects designs [32]. Specifically, the estimation of treatment \times subject interactions, or the size of the effect of training on counselor skills in this case, is direct, whereas the effect of training on individual counselor skills is masked by sampling error in between-subjects designs. Hunter & Schmidt argue that the power to detect the true effect of treatment on the treated is higher in case-control studies than it is in experimental designs and argue that discarding case-control designs without consideration of their strengths compared to the between-subjects designs is unwarranted.

Means and standard deviations of each measurement occasion were coded for each observational outcome in each study. Reliability coefficients were also coded for the main outcome variables. The calculation and interpretation of effects sizes and their associated sampling error variances followed Lipsey & Wilson [33].

Change scores were calculated for each experimental and control condition in each study from pre-training to post-training and from post-training to the 3, 6 and 6+ months follow-up periods. Therefore, pre-post change scores are interpreted as the gain in MI skills attributable to the workshop, and post-training change scores are interpreted as the gain or deterioration of MI skills following participation in the training workshop (change score less than 0.20 = small, change score between 0.20 and 0.80 = medium, change score greater than 0.80 = large [34]). Preliminary analysis failed to identify a moderating effect of outcome measure [i.e. Motivational Interviewing Treatment Integrity (MITI), Motivational Interviewing Skills Code (MISC), other] on training outcomes. Therefore, effect sizes for all observational measures were averaged within each study condition to produce a single effect size for each study condition.

Analysis

Hunter & Schmidt random-effects models were employed to calculate inverse-variance weighted effect sizes, correct

standard errors for sampling error variance and to calculate credibility intervals [32]. Effect sizes were corrected to account for small sample size bias using the formula $d_a = d / (1 + [3 / (4n - 12)])$. Categorical moderators were tested by estimating average effect sizes and credibility intervals for each category. Moderators on continuous scales were categorized into groups based on the sample distribution described below. Missing data was treated using listwise deletion. All analyses were conducted manually using formulas programmed in Excel (worksheets available upon request).

RESULTS

Sample characteristics

The 21 studies in this meta-analysis involved a total of 987 service providers in 37 experimental or control groups (see Table 1). Half the studies employed random assignment, while the remaining half utilized quasi-experimental and case-control designs. Most studies (62%) trained practitioners in substance abuse treatment settings, while the remaining studies (38%) involved health professionals and correction officers. More than half the studies (62%) used recorded interviews with actual patients to measure trainee skills. Thirteen studies (62%) included a 'workshop only' training condition, while 13 studies (62%) included post-workshop training inputs.

Twelve studies utilized the MITI code [31]. Three behavioral outcome measures were widely reported, including MI adherence, MI spirit and reflection : question ratio (R : Q). Studies that used the MITI reported on average 2.27 behavioral outcome measures; 81% reported MI consistent, 73% reported MI spirit and 73% reported the R : Q ratio. Six studies utilized the MISC [30]. The same three behavioral measures (or their equivalents) were reported, in addition to the percentage of open-ended measures, and percentage of complex reflections. Studies using the MISC reported an average 2.5 behavioral outcome measures. All studies that used MISC reported MI consistent, 67% reported MI spirit and 83% reported R : Q. Two studies utilized the Independent Tape Rater Scale (ITRS), and both reported behavioral outcome measures of adherence and competence [22,23]. One study utilized Peer Proficiency Assessment (PEPA), which did not include behavioral measures of MI adherence, MI spirit or R : Q [24].

Table 2 presents the main findings of the study. MI training workshops, overall, yielded an effect size of 0.76 at post-workshop assessment ($k = 23$) compared to the more modest and non-significant gains for trainees in 'self-study' conditions [$k = 3$, $d = 0.56$, 95% confidence interval (CI) = -0.39 to 1.52, $I^2 = 94.71$]. The overall

Table 2 Post-training and follow-up effect sizes for workshop only and post-training enhancement conditions.

	<i>k</i>	<i>d</i>	95% CI ^a	<i>I</i> ²
Any active training				
Post-test	23	0.76	0.43 to 1.10	50.76
3 months	13	-0.17	-0.66 to 0.32	78.81
6 months	17	-0.04	0.32 to 0.24	47.98
Standard workshop				
Post-test	12	0.78	0.43 to 1.12	67.07
3 months	8	-0.11	-0.69 to 0.46	86.06
6 months	8	-0.01	-0.33 to 0.31	44.88
Workshop enhancements				
Post-test	11	0.75	^b	7.19
3 months	5	-0.22	-0.42 to 0.00	53.68
6 months	9	-0.03	^b	0.00
Workshop only				
Post-test	11	0.76	0.42 to 1.11	58.10
3 months	5	-0.35	^b	20.80
6 months	4	-0.30	^b	25.71
Any sustaining				
Post-test	12	0.75	0.28 to 1.22	64.10
3 months	8	0.04	-0.36 to 0.44	69.33
6 months	13	0.03	-0.14 to 0.20	36.15
Feedback and coaching				
Post-test	9	0.75	0.19 to 1.31	70.89
3 months	7	-0.03	-0.45 to 39	72.44
6 months	8	0.04	^b	16.11
Coaching only				
Post-test	2 ^c	-	-	-
3 months	0 ^c	-	-	-
6 months	3	-0.08	^b	0

Post-test change scores indicate change from pre-training to post-workshop on behavioral measures of motivational interviewing (MI) skills; 3- and 6-month effect sizes indicate change following the post-workshop measures of MI skills. ^aCredibility intervals corrected for sampling error variance [32]. ^bVariance was fully attributed by sampling error variance. Credibility intervals cannot be calculated. ^cStudies were not meta-analysed when there were fewer than three contributing samples. CI = confidence interval.

effects of training began to erode by the 3-month follow-up ($d = -0.17$) and recovered somewhat by the 6-month follow-up ($k = 18$, $d = -0.04$). These effect sizes were characterized by high levels of heterogeneity, with I^2 statistics ranging from 48 to 79%. Thus, the following analyses sought to determine the source of the heterogeneity in three areas: workshop enhancements, structure of post-workshop training enhancements and methodological moderators.

Workshop enhancements

The first analysis examined the moderating effect of workshop design (see Table 2). Workshops generally lasted for 12–15 hours, conducted over the course of 2 days ($k = 23$), and consisted of face-to-face didactic

instruction and interactive exercises. Workshop enhancements included dividing the workshop into smaller segments in a context-tailored approach ($k = 3$), use of booster sessions ($k = 3$), providing feedback to trainees using audio-recorded practice samples ($k = 8$), the use of train-the-trainer models ($k = 1$) and the use of video-, web- and computer-based technologies to administer the training ($k = 2$). At post-workshop assessments there were no differences between standard workshops ($d = 0.78$) and workshops that included some innovative element (i.e. context-tailored, booster session, feedback, train the trainer, web) ($d = 0.75$). The type of workshop enhancement was not associated with effect size. Workshop enhancements were also not associated with long-term sustainability. Aggregating across all follow-up periods, change scores reflected eroding skills for studies that reported workshop enhancements without post-workshop training activities ($k = 6$, $d = -0.27$, 95% CI = -0.65 to 0.12 , $I^2 = 85.90$).

Post-workshop inputs

The second analysis examined the moderating effect of post-workshop inputs on the sustainability of MI skills. Table 2 summarizes these results. Effect sizes among workshop-only studies evidenced eroding skills by the 3- ($d = -0.35$) and 6-month follow-up periods ($d = -0.30$). Post-workshop inputs increased retention of motivational interviewing skills. Non-significant change scores at 3 ($d = 0.04$) and 6 months ($d = 0.03$) suggest that post-workshop training inputs sustained MI skills over a 6-month time-frame. It appears that by 6 months, feedback and coaching together ($d = 0.04$) is slightly more effective than coaching only ($d = -0.08$).

Table 3 shows that follow-up effect sizes varied according to the structure of post-workshop feedback and coaching activities. Specifically, increasing the frequency of trainings, spreading the training out across 6 months and increasing the number of hours of post-workshop training seem to yield more enduring effect sizes. By 6 months, training programs that included three to four post-workshop contacts yielded a combined effect size of 0.01, and training programs that included five to eight post-workshop contacts yielded a combined effect size of 0.06. Expanding the time-frame during which training occurs was shown to enhance skill retention. By 6 months, trainings that lasted 10–16 weeks yielded a combined effect size of -0.03 , and trainings that lasted 20–24 weeks yielded a combined effect size of 0.11. In terms of hours of post-workshop contacts, Table 3 shows that studies with post-training contact hours of 5–12 hours or greater had increased skills at the 6-month follow-up ($d = 0.19$) compared to studies with fewer than 5 hours of post-training contact hours, which evidenced a slight decline over this period ($d = -0.06$).

Table 3 Frequency, duration and intensity of post-workshop feedback and coaching.

	<i>k</i>	<i>d</i>	95% CI	<i>I</i> ²
Contacts				
3–4 contacts				
3 months	2 ^a	–	–	–
6 months	8	0.01	^b	0
5–8 contacts				
3 months	4	0.09	–0.40 to 0.57	77.36
6 months	4	0.06	–0.33 to 0.39	73.34
Weeks				
10–16 weeks				
3 months	6	–0.01	–0.29 to 0.26	62.92
6 months	5	–0.03	–0.12 to 0.07	34.38
20–24 weeks				
3 months	0 ^a	–	–	–
6 months	4	0.11	^b	0
52 weeks+				
3 months	1 ^a	–	–	–
6 months	4	–0.00	–0.20 to 0.19	43.93
Hours				
Below median (<5 hours)				
3 months	3	–0.02	–0.32 to 0.28	72.89
6 months	7	–0.06	^b	4.03
Above median (≥5 hours)				
3 months	5	0.05	–0.50 to 0.60	76.37
6 months	6	0.19	^b	0

^aStudies were not meta-analysed when there were fewer than three contributing samples. ^bVariance was fully attributed by sampling error variance. Credibility intervals cannot be calculated. CI = confidence interval.

Methodological moderators

The effect of eight additional methodological moderators were examined: setting type (SUD versus other, agency-based versus training center-based), training adaptations for the agency context, trainee education level (percentage with masters' degree or higher), research design (experimental versus case-control), outcome measure (MITI/MISC versus other), sample size and attrition rate. From this list, only three moderated effect sizes: education, sample size and attrition.

The moderating effect of education was statistically significant, although small. Using a median split for percentage of the study sample with a masters' degree or higher (median = 64%), results indicated that samples with education levels below the median had slightly higher 6-month effect sizes than studies on samples with higher levels of education ($k = 10$, $d = 0.07$ versus $d = -0.05$, $P < 0.05$). Despite the statistically significant difference, the effect size for both groups suggests relatively stable outcomes over the follow-up period.

The moderating effect of sample size was slightly stronger. Although results were equivocal when using a median split to compare small samples with larger samples, the pattern became clear when we compared the

five largest samples with the five smallest samples. The five studies with the smallest sample sizes (mean = 10.0) had a larger effect size than the five studies with the largest sample sizes (mean = 27.6) ($d = 0.18$ versus $d = -0.03$, $P < 0.05$).

Attrition rates had a larger moderating impact on effect sizes. Attrition rates were calculated for each study by coding the total possible number of post-workshop training inputs, dividing the total possible by the average number of post-workshop training inputs received, and subtracting this quantity from one. Twelve studies with post-workshop inputs provided sufficient information to calculate attrition rates. Attrition ranged from 0 to 86% (mean_{weighted} = 32%; median = 29%). Samples with attrition below the median had an average increase in MI skills over a 6-month follow-up period ($k = 7$, $d = 0.12$, 95% CI = 0.06–0.18, $I^2 = 0$), whereas samples with attrition above the median had an average decrease in MI skills over the same follow-up period ($k = 4$, $d = -0.10$, $I^2 = 25.36$). Two studies were dropped from this analysis because their longest follow-up period was 3 months. When the analysis was replicated with the longest follow-up period available for each study with attrition above the median, the moderating effect of attrition became stronger ($k = 6$, $d = -0.29$, $I^2 = 0$).

DISCUSSION

This study examined training approaches associated with the long-term stability of MI skills for human service workers in as-usual treatment settings. Overall, the results of the study quantify the conventional wisdom of the field—that skills developed during the course of traditional training workshops erode quickly when additional post-workshop training inputs are not provided. This finding accords with the general training literature on behavioral interventions that support the use of multimodal training approaches [1]. Across evidence-based interventions, neither self-directed training nor workshop-based training has been associated with sustained skillfulness in trainees without the addition of post-training supervision. The results of this study suggest that the level of post-training expert supervision needed to sustain MI skills is somewhat modest—approximately three to four contacts totaling at least 5 hours of contact time over a 6-month period was sufficient for the average study to sustain training effects over a 6-month window.

The moderator analyses involving sample size and attrition seemed to converge with this overall pattern of results. When taken as proxies for intensity of contact between trainer and trainee, training workshop effects were sustained when trainees had more contact with their trainers and expert supervisors. Indeed, under more optimal conditions with small training group sizes and

low attrition, further skill gains were realized with the addition of post-workshop feedback and coaching.

These findings point to an opportunity to conduct further research on the nature of post-workshop training inputs. Promising innovations include train-the-trainer models [22] and *in-vivo* supervision, where instantaneous feedback is provided to practitioners during live treatment sessions [6]. Absent from the literature were feedback methods that did not include the coding of actual treatment samples. To the extent that the current feedback approach remains resource-intensive and potentially unacceptable to at least some treatment providers, attrition from training protocols that include feedback on specific audio-recorded sessions may limit the reach of MI training.

The data in this study suggest that the workshop structure itself does not moderate long-term training outcomes. Quite possibly, heterogeneity in the category we called 'workshop enhancements' masked the effect of some types of workshop characteristics on long-term training outcomes. Moreover, it is likely that published descriptions of training workshops do not include the full range of enhancements that are frequently employed by skilled trainers. Nevertheless, the lack of moderation may also reflect the overall finding of this analysis, that the critical ingredient for long-term sustainability of MI skillfulness among trainees is adherence to a modest level of ongoing coaching, feedback and supervision. Exploration of this finding awaits further research.

It should be noted that limitations in the sample of studies available for this review temper conclusions that can be drawn from this study. First, although the findings of this review suggest that a modest dose of post-workshop supports is needed to sustain MI skillfulness over a 6-month follow-up period, widespread application of this finding in practice should await confirmation in more rigorous experimental research. With limited exceptions [2,6], there is a dearth of rigorous research that experimentally manipulates the type and intensity of post-workshop supervision provided to trainees. We accommodated this limitation through the use of a within-subjects design rather than the preferred between-subjects design. Secondly, the limited sample size did not allow us to examine simultaneously the independent effect of the post-workshop moderators. Therefore, it is not clear which post-workshop dose measure, number of contacts, weeks of post-workshop inputs or contact hours was associated with more or less skill retention, for example. Thirdly, there were too few studies with follow-up periods greater than 6 months to make inferences about the erosion of MI skills once post-workshop enhancements are withdrawn. Fourthly, there were too few studies reported in non-SUD treatment

agencies to conduct a detailed analysis of training adaptations needed to foster sustained MI skillfulness. Finally, the pool of study reports did not uniformly present characteristics about trainee characteristics (e.g. educational level). More detailed information about trainee characteristics is needed to optimize training approaches.

Interest in MI and its widespread implementation has outpaced the literature on MI effectiveness. One area of ongoing research that bears on the implementation of this review in direct practice is the isolation of active ingredients and mechanisms of change. This nascent literature has examined the effect of specific MI components commonly measured with the MITI and MISC, including global measures of MI spirit, MI-consistent counselor behaviors and MI-inconsistent counselor behaviors, along with specific counselor techniques such as complex reflections, open-ended questions and affirmations. To date, the data have failed to coalesce around specific MI techniques.

The results of this study have direct implications for agency practice and research. First, it appears that radical shifts in supervisory practices may not be required to sustain staff skillfulness in many settings. Rather, the addition of approximately monthly coaching and feedback sessions devoted to MI implementation and skills can play an important role in the sustainability of MI skills. Secondly, it appears that skill gains are possible over post-workshop supervisory periods when trainees maintain their involvement in the feedback and coaching process. While most training programs place their emphasis on workshops as the primary vehicle for learning, the results of this study hint at the importance of post-training supports. These supports need not be limited to sustaining training gains but may, instead, actively advance trainees towards overall MI proficiency. Finally, new methods for monitoring the practice skills of practitioners are needed. Current methods, relying as they do on invasive recording and expensive coding procedures, are useful for the process of coaching and feedback but are not sustainable for many agencies.

References

1. Herschell A. D., Kolko D. J., Baumann B. L., Davis A. The role of therapist training in the implementation of psychosocial treatments: a review and critique with recommendations. *Clin Psychol Rev* 2010; 30: 448–66.
2. Miller W. R., Yahne C. E., Moyers T. B., Martinez J., Pirritano M. A randomized trial of methods to help clinicians learn motivational interviewing. *J Consult Clin Psychol* 2004; 72: 1050–62.
3. Barwick M. A., Bennett L. M., Johnson S. N., McGowan J., Moore J. E. Training health and mental health professionals in motivational interviewing: a systematic review. *Child Youth Serv Rev* 2012; 34: 1786–95.

4. Madson M. B., Loignon A. C., Lane C. Training in motivational interviewing: a systematic review. *J Subst Abuse Treat* 2009; **36**: 101–9.
5. de Roten Y., Zimmermann G., Ortega D., Delpland J. Meta-analysis of the effects of MI training on clinicians' behavior. *J Subst Abuse Treat* 2013; **45**: 155–62.
6. Smith J. L., Carpenter K. M., Amrhein P. C., Brooks A. C., Levin D., Schreiber E. A. *et al.* Training substance abuse clinicians in motivational interviewing using live supervision via teleconferencing. *J Consult Clin Psychol* 2012; **80**: 450–64.
7. Moyers T. B., Manuel J. K., Wilson P. G., Hendrickson S. M., Talcott W., Durand P. A randomized trial investigating training in motivational interviewing for behavioral health providers. *Behav Cogn Psychother* 2008; **36**: 149–62.
8. Aarons G. A., Hurlburt M., Horwitz S. M. Advancing a conceptual model of evidence-based practice implementation in public service sectors. *Adm Policy Ment Health* 2011; **38**: 4–23.
9. Fixsen D. L., Blase K. A., Naoom S. F., Wallace F. Core implementation components. *Res Soc Work Pract* 2009; **19**: 531–40.
10. Greenhalg T., Robert G., MacFarlane F., Bate P., Kyriakidou O. Diffusion of innovations in service organizations: systematic review and recommendations. *Milbank Q* 2004; **82**: 581–629.
11. Baer J. S., Wells E. A., Rosengren D. B., Hartzler B., Beadnell B., Dunn C. W. Agency context and tailored training in technology transfer: a pilot evaluation of motivational interviewing training for community counselors. *J Subst Abuse Treat* 2009; **37**: 191–202.
12. Moher D., Liberati A., Tetzlaff J., Altman D. G., TP Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLOS Med* [internet] 2009; **6**: e1000097. Available at: <http://www.plosmedicine.org> (accessed 26 June 2012) (Archived at <http://www.webcitation.org/6PCX27Aie> on 29 April 2014).
13. Baer J. S., Rosengren D. B., Dunn C. W., Wells E. A., Ogle R. L., Hartzler B. An evaluation of workshop training in motivational interviewing for addiction and mental health clinicians. *Drug Alcohol Depend* 2004; **73**: 99–106.
14. Bennett G. A., Moore J., Vaughan T., Rouse L., Gibbins J. A., Thomas P. *et al.* Strengthening motivational interviewing skills following initial training: a randomised trial of workplace-based reflective practice. *Addict Behav* 2007; **32**: 2963–75.
15. Bohman B., Forsberg L., Ghaderi A., Rasmussen F. An evaluation of training in motivational interviewing for nurses in child health services. *Behav Cogn Psychother* 2013; **41**: 329–43.
16. Britt E., Blampied N. M. Motivational interviewing training: a pilot study of the effects on practitioner and patient behaviour. *Behav Cogn Psychother* 2010; **38**: 239–44.
17. Brug J., Spikmans F., Aartsen C., Breedveld B., Bes R., Fereira I. Training dietitians in basic motivational interviewing skills results in changes in their counseling style and lower saturated fat intakes in their patients. *J Nutr Educ Behav* 2007; **39**: 8–12.
18. Evangeli M., Engelbrecht S., Swartz L., Turner K., Forsberg L., Soka N. An evaluation of a brief motivational interviewing training course for HIV/AIDS counsellors in Western Cape Province, South Africa. *AIDS Care* 2009; **21**: 189–96.
19. Evangeli M., Longley M., Swartz L. No reductions and some improvements in South African lay HIV/AIDS counsellors' motivational interviewing competence one year after brief training. *AIDS Care* 2011; **23**: 269–73.
20. Forsberg L., Forsberg L. G., Lindqvist H., Helgason A. R. Clinician acquisition and retention of motivational interviewing skills: a two-and-a-half-year exploratory study. *Subst Abuse Treat Prev Policy* [internet] 2010; **5**: 8.
21. Hartzler B., Baer J. S., Dunn C. W., Rosengren D. B., Wells E. A. What is seen through the looking glass: the impact of training on practitioner self-rating of motivational interviewing skills. *Behav Cogn Psychother* 2007; **35**: 431–45.
22. Martino S., Ball S. B., Nich C., Canning-Ball M., Rounsaville B. J., Carroll K. M. Teaching community program clinicians motivational interviewing using expert and train-the-trainer strategies. *Addiction* 2011; **106**: 428–41.
23. Martino S., Canning-Ball M., Carroll K. M., Rounsaville B. J. A criterion-based stepwise approach for training counselors in motivational interviewing. *J Subst Abuse Treat* 2011; **30**: 357–65.
24. Mastroleo N. R., Turrissi R., Carney J. V., Ray A. E., Larimer M. E. Examination of posttraining supervision of peer counselors in a motivational enhancement intervention to reduce drinking in a sample of heavy-drinking college students. *J Subst Abuse Treat* 2010; **39**: 289–97.
25. Miller W. R., Mount K. A. A small study of training in motivational interviewing: does one workshop change clinical and client behavior? *Behav Cogn Psychother* 2001; **29**: 457–71.
26. Shafer M. S., Rhode R., Chong J. Using distance education to promote the transfer of motivational interviewing skills among behavioral health professionals. *J Subst Abuse Treat* 2004; **26**: 141–8.
27. Smith J. L., Amrhein P. C., Brooks A. C., Carpenter K. M., Levin D., Schreiber E. A. *et al.* Providing live supervision via teleconferencing improves acquisition of motivational interviewing skills after workshop attendance. *Am J Drug Alcohol Abuse* 2007; **33**: 163–8.
28. van Eijk-Hustings Y. J. L., Daemen L., Schaper N. C., Vrijhoef J. J. M. Implementation of motivational interviewing in a diabetes care management initiative in the Netherlands. *Patient Educ Couns* 2010; **84**: 10–5.
29. Walters S. T., Vader A. M., Nguyen N., Harris T. R., Eells J. Motivational interviewing as a supervision strategy in probation: a randomized effectiveness trial. *J Offender Rehabil* 2010; **49**: 309–23.
30. Miller W. R., Moyers T. B., Ernst D., Amrhein P. Manual for the Motivational Interviewing Skill Code (MISC). Albuquerque, NM: Center on Alcoholism, Substance Abuse, and Addictions, the University of New Mexico; 2008.
31. Moyers T. B., Manuel J. K., Miller W. R., Ernst D. Revised global scales: Motivational Interviewing Treatment Integrity 3.1.1 (MITI 3.1.1). Albuquerque, NM: University of New Mexico, Center on Alcoholism, Substance Abuse and Addictions (CASAA); 2010.
32. Hunter J. E., Schmidt F. *Methods of Meta-Analysis*, 2nd edn. Thousand Oaks, CA: Sage; 2004.
33. Lipsey M. W., Wilson D. B. *Practical Meta-Analysis*. Thousand Oaks, CA: Sage Publications; 2001.
34. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*, 2nd edn. Hillsdale, NH: Lawrence Erlbaum Associates; 1988.