January 2022

A Systematic Review And Meta-Analysis Of Motivational Interviewing Training Effectiveness Among Students-In-Training

Amy K. Maslowski

[How does access to this work benefit you? Let us know!](#)

Follow this and additional works at: https://commons.und.edu/theses

Recommended Citation

A SYSTEMATIC REVIEW AND META-ANALYSIS OF MOTIVATIONAL
INTERVIEWING TRAINING EFFECTIVNESS AMONG STUDENTS-IN-TRAINING

by

Amy Kristine Maslowski

Bachelor of Arts, College of Saint Benedict, 2016
Master of Arts, University of Minnesota Duluth, 2018

A Dissertation

Submitted to the Graduate Faculty

of the

University of North Dakota

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Grand Forks, North Dakota

August,
2022

Name: Amy K. Maslowski

Degree: Doctor of Philosophy

This document, submitted in partial fulfillment of the requirements for the degree from the University of North Dakota, has been read by the Faculty Advisory Committee under whom the work has been done and is hereby approved.

Cindy L. Juntunen, Ph.D.

Rhea L. Owens, Ph.D.

Klaus E. Cavalhieri, Ph.D.

Virginia Clinton-Lisell, Ph.D.

Rick A. LaCaille Ph.D.

This document is being submitted by the appointed advisory committee as having met all the requirements of the School of Graduate Studies at the University of North Dakota and is hereby approved.

Chris Nelson
Dean of the School of Graduate Studies

11/2/2020

Date

iii

PERMISSION

Title	A Systematic Review and Meta-Analysis of Motivational Interviewing Training Effectiveness Among Students-in-Training
Department	Counseling Psychology
Degree	Doctor of Philosophy

In presenting this dissertation in partial fulfillment of the requirements for a graduate degree from the University of North Dakota, I agree that the library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the professor who supervised my dissertation work or, in her absence, by the Chairperson of the department or the dean of the School of Graduate Studies. It is understood that any copying or publication or other use of this dissertation or part thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of North Dakota in any scholarly use which may be made of any material in my dissertation.

Amy K. Maslowski
October 8, 2020

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure

ABSTRACT

Effectively using motivational interviewing (MI) in practice can be difficult. However, there are a number of studies that examine training students across helping professions with the goal of facilitating students use MI more effectively. Although there is no standardized training manual, students often learn specific MI skills (e.g., open-ended questions, reflections) and knowledge (e.g., MI spirit) in hopes that they will apply those techniques to encounters with clients. The purpose of this systematic review and meta-analysis was to quantify the effectiveness of teaching students motivational interviewing. In total, 15 randomized and non-randomized studies met inclusion criteria and were examined in the current review of 8 dependent variables. A large and significant aggregated Hedges' g of 0.90 (95% CI [0.45, 1.35]) was found. However, large heterogeneity was observed in all but one of the dependent variables. Moderation analyses revealed no significant moderating effects for risk of bias or type of comparison group; however, training length was a significant moderator. Limitations of the current meta-analysis include the small sample size and lack of consistency among training duration, measurement, and data collection and resulting heterogeneity. Future research appears warranted to further assess student MI training effectiveness, especially using more rigorous and standardized procedures, as well as determining enduring effects of the training.

Keywords: motivational interviewing, MI, training, students, meta-analysis

A Systematic Review and Meta-Analysis of Motivational Interviewing Training

Effectiveness Among Students-in-Training

Motivational interviewing (MI) is a common therapeutic approach used by both mental health and health professionals (e.g., Arkowitz et al., 2015; Miller & Rollnick, 2013; Rollnick et al., 2008). Previous research (e.g., Lundahl et al., 2010; Schwalbe et al., 2014) has supported the effectiveness of training providers in these fields. More recently, MI has been incorporated into students' training curriculum within their respective fields (e.g., medical students; Daeppen et al., 2012; counseling students; Young & Hagedorn, 2012). Considerably less research, however, has been conducted to investigate the effectiveness of training students in MI. The purpose of this study was to systematically review and quantify the effects of teaching students-in-training motivational interviewing skills.

Motivational Interviewing

MI was created by Miller and Rollnick (1992) as an approach to assist professionals in collaboratively working with clients to promote change with substance abuse difficulties (Miller & Rollnick, 2013; Rollnick et al., 2008). Briefly, MI involves *rolling with resistance* – working with clients to elicit their motivation to change – when ambivalence is evident. Furthermore, MI is directive, but also client-centered. MI can be the critical bridge between "wrestling" with a client (figuratively) and "dancing" with them to make a change (Rollnick et al., 2008). The spirit of MI promotes autonomy, collaboration, and evocation (ACE), such that: (a) MI is *autonomy*-supportive, meaning it allows the client choice in the change process; (b) MI endorses a *collaborative* nature between the client and provider in working towards the client's goals and change; and (c) MI providers strive to *evoke* motivation and reasons for changing directly from the client (Miller & Rollnick, 2013).

MI with Mental & Behavioral Health Conditions

Overall, research largely supports the use of MI with clients who have mental health conditions (Arkowitz et al., 2015), including addiction concerns (including gambling; DiClemente et al., 2017) as well as comorbid conditions (Rollnick et al., 2008). Meta-analyses and systematic reviews have found support for the use of MI with eating disorders (Macdonald et al., 2012), severe mental illness and comorbid substance use (Cleary et al., 2008), diabetes (when targeting one behavior at a time; Hunt et al., 2013), and, generally, to increase medication adherence (Palacio et al., 2016). One known study (Vanderwall, 2015) found mixed results for the use of MI with people with schizophrenia; MI was not suggested as the first line of treatment with this population. Thus, although reviews have generally found positive efficacy for MI with various mental health concerns, researchers suggest additional studies should be conducted to further investigate their relationship and outcomes with specific conditions, in addition to how well MI works as a standalone treatment compared to integrated with another approach (e.g., MI + cognitive-behavioral therapy; CBT; Iarussi, 2019).

MI is also one of the most common interventions used in integrated care settings to effectively address behavioral and mental health conditions in brief interactions and is widely used by behavioral health consultants (BHCs; Dunn et al., 2001; Moyers et al., 2007). BHCs trained in MI generally assist clients with substance use and/or increasing physical activity (Dunn et al., 2001; Moyers et al., 2007). They work to integrate components of both physical and mental health to promote positive behavior change. Clients are usually seen by BHCs as part of their medical visit; thus, their time with clients is limited. Therefore, BHCs use, for example, open-ended questions and reflections to help develop discrepancies between the client's behavior and well-being and often assign homework. The use of MI by behavioral health consultants has

been shown to be associated with a decrease in clients' substance use/consumption, cigarette use, and risk-taking behaviors (e.g., unprotected sexual intercourse), as well as an increase in clients' physical activity, and, generally, with treatment adherence (Moyers et al., 2007).

MI with Health Conditions

Medical professionals generally have a short time to interact with clients, so they need to make the most of these exchanges. MI offers an efficient and effective approach to doing so in the medical setting (Rollnick et al., 2008). MI has been shown to be effective with clients who have concerns such as weight loss, unhealthy eating habits, and smoking and alcohol use, in addition to other concerns, including asthma and chronic obstructive pulmonary disease (COPD), cardiovascular health, odonatological needs, diabetes, difficulties with medical and/or treatment adherence, and pain (Rollnick et al., 2008).

Lundahl and colleagues (2013) recently conducted a meta-analysis and systematic review of the effectiveness of MI with several health conditions. Their sample included 48 randomized controlled trials (RCTs) with 51 independent effect sizes and 9,618 clients. A significant and positive odds ratio (OR) of 1.55 was found; however, heterogeneity was high ($I^2 = 90.42$; Lundahl et al., 2013), suggesting that there was significant variability in the operationalization of outcomes among the included studies. Moderation analyses revealed that MI effectiveness did vary by problem types, such that MI had positive effects for clients who struggled with blood pressure, cholesterol, HIV viral load, and medical/treatment adherence. Furthermore, use of MI helped clients decrease alcohol, tobacco, and marijuana usage. Results were more mixed with healthy living behaviors (e.g., weight management, eating), risk-related behaviors (e.g., safe sex), and quality of life (e.g., independence). Mental health professionals and mixed teams were the only provider types with significant moderating effects on outcomes. Mental health

professionals had a slightly larger effect (OR = 1.73) than mixed teams (OR = 1.23), though these were not significantly different. Notably, MI was also found to be durable up to 13 months post-treatment. Overall, this study indicated that MI is efficacious with several health conditions and in a myriad of settings (Lundahl et al., 2013). Collectively, these studies provide evidence for MI's effectiveness to be used in a short period of time and with providers who have less background in mental health.

Training in MI

The framework often used for training professionals to use MI has been conceptualized into eight stages (Miller & Moyers, 2006). The stages are sequential and build upon previously acquired skills. The stages include learning: (a) the *spirit of MI* (i.e., ACE); (b) *client-centered counseling skills* (i.e., open-ended questions, affirmations, reflections, summaries; OARS); (c) how to *recognize and reinforce change talk* (i.e., client language that suggests moving toward change; Miller & Rollnick, 2013); (d) how to *elicit and strengthen change talk;* (e) how to *roll with resistance* (i.e., responding to the client's reasons not to change without contest often through reflections; Miller & Rollnick, 2013); (f) how to *develop a change plan*, or goals for next steps*;* (g) how to consolidate the *client's commitment* to making changes and the change process; and (h) how to *switch between MI and other counseling methods,* as well as integrating MI with other approaches (Miller & Moyers, 2006). Notably, *resistance* has recently been replaced by two new terms: *discord* (i.e., tension or defensiveness in response to the therapist's negative judgement and/or control; Miller & Rollnick, 2013) and *sustain talk* (i.e., the client's verbalized ambivalence about behavior change), though the therapist still responds in the same way (Miller & Rollnick, 2013).

Previous Reviews of Training Studies

Several previous reviews have both systematically (Barwick et al., 2012; Hall et al., 2016; Madson & Campbell, 2006; Madson et al., 2009; Söderlund et al., 2011) and meta-analytically (de Roten et al., 2013; Lundahl et al., 2010; Schwalbe et al., 2014) reviewed the effectiveness of MI training studies with various professionals. Overall, these reviews provided positive support for increasing professionals' helping (e.g., open-ended questions, empathy) and MI-specific skills (e.g., reflections).

Notably, Lundahl and colleagues' (2010) meta-analysis focused on professional trainees, but it also included some student trainee studies. The findings from this study suggested that one's training background (e.g., nursing, psychology) does not influence the success of the client encounter, provided the trainee effectively empathizes with others. Nonetheless, they found student trainee moderation effects. For example, when MI was contrasted to comparison groups (i.e., waitlist, written materials only, non-specific treatment-as-usual) and delivered by mental health clinicians with a Master's or Ph.D., it was more effective (Hedges' $g = 0.39$) than delivery by students still in school ($g = 0.23$), mental health providers with a bachelor's degree ($g = 0.19$), or nurses ($g = 0.10$). Lastly, study quality/rigor was assessed as a potential moderating variable and was only found to be significant when comparing MI to strong comparison groups (e.g., specific treatment-as-usual); higher study quality was associated with smaller effect sizes.

Previous Reviews Limitations

Research investigating the effects of MI training among students and professionals has garnered generally positive support for its implementation within the trainee populations examined. One significant limitation, however, involves inconsistent skills assessment (i.e., skills and/or behaviors being assessed without a standardized measure and, thus, with varied reliability

and validity). Some studies used validated measures (e.g., Motivational Interviewing Treatment Integrity; MITI; Moyers et al., 2005), whereas others assessed skill use through trainee self-report. Self-report data comes with response bias concerns (e.g., social desirability, response-shift bias; Rosenman et al., 2011). Furthermore, existing training studies with professionals point to the gap of lacking follow-up and/or supervision sessions after the initial trainings; these would likely help to produce more enduring effects for the trainees (e.g., de Roten et al., 2013; Lundahl et al., 2010). Moreover, many studies either used only a short follow-up period (e.g., one month) or did not follow-up at all.

Finally, study quality observed in previous reviews has varied considerably. The only meta-analysis that assessed the potential moderating effects of study quality was Lundahl and colleagues' (2010) study. Several other MI training meta-analyses and systematic reviews with professionals had considerable variability in study quality (Barwick et al., 2012; de Roten et al., 2013; Söderlund et al., 2011); however, the variable was not assessed as a moderator. The remaining previously published MI training reviews (Hall et al., 2016; Madson & Campbell, 2006; Madson et al., 2009; Schwalbe et al., 2014) did not discuss study quality. Study quality is an important variable to assess because of the potentially confounding effects (e.g., selection bias, attrition bias; Ryan et al., 2013) it can have on study findings and, thus, meta-analysis and systematic review results.

Student Trainings

Students-in-training (e.g., mental health, medical), like professionals, face situations where clients present with a myriad of concerns. Students in mental health-related fields (e.g., counseling, social work) may assist clients with lifestyle adaptations (e.g., weight loss, medication adherence), as well as modifications involving their well-being (e.g., engagement

with supportive people). For medical students, presenting concerns may range from treatment adherence to poor physical activity and nutrition. The common thread among both professionals and students is that they are commonly trying to promote client behavior change.

Although MI is being taught to students in both mental-health and health-related programs, results across studies on the efficacy of MI training programs is mixed and the outcomes measured are inconsistent. Importantly, no standardized MI training or structured protocol is available exclusively for students. Few student-only studies have used the full two-day standard MI training (Hodorowicz, 2018; Pecukonis et al., 2016). Whereas some studies have used brief MI training (BMI; e.g., Han & Cho, 2018; Young & Hagedorn, 2012), though the structure and content of the trainings varied considerably among these studies. Much of the previous literature involving training students in MI has focused on teaching MI-consistent skills (e.g., empathy, open-ended questions). However, unstandardized outcome assessment and varying length of follow-up continue to persist as limitations in student training research.

The only known review involving MI student training studies was conducted by Dunhill and colleagues (2014) with medical students. The studies that were part of this review varied in the type of MI training provided to students; some used a workshop format while others included a MI rotation. Overall, the researchers found positive effects from the training, including students improving their MI skills and competency, and an increased favorable attitude towards using MI. However, the lack of control groups within the included studies, along with a limited use of validated outcome measures in the studies reviewed, reduces confidence in the conclusions one can draw from this review (Dunhill et al., 2014).

Objectives

Overall, previous research involving training students in MI varies considerably in both purpose and overall effectiveness. Therefore, the aim of the current meta-analysis and systematic review was to provide the first investigation into the effectiveness of MI trainings for students – regardless of healthcare specialty area – including those who are training in professional capacities (e.g., medical school). Three potential moderators were explored, including study quality, type of comparison group, and length of training.

The following research questions were addressed:

1. What is the effectiveness of MI training with students-in-training?

2. What study variables (i.e., risk of bias, training length) moderate the effects of MI training with students? Furthermore, given limited literature examining student trainees, one additional moderator (i.e., type of comparison group) was explored, as previous research indicates the type of comparison group used (e.g., another active treatment, waitlist control) have different active components, which may influence the results (Karlsson & Bergmark, 2015).

Method

A review protocol, including a methodology and analysis plan, was published prior to conducting the current study (Maslowski, 2019). The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al., 2009) were also used in the current review (see Appendix A).

Search Strategy & Study Selection

A comprehensive literature search was conducted in May, 2019 to gather all relevant empirical studies meeting the designated inclusion criteria. One additional search was conducted

in September, 2019 to check for any new studies. These included searches in: (a) online databases for both published studies and dissertations or theses, (b) online databases within journals that commonly publish research in this area, (c) reference sections of included studies and studies of interest, (d) the MINT website, and (e) more recent studies that cited previously published studies found from approaches (a-d) through the Google Scholars' "cited" feature. Search terms used included: *motivational interviewing*, *MI*, *training*, *education, student*, trainee*, intern*, and resident** from the following databases: PsycINFO, PsycARTICLES, Consumer Health Complete, PubMed, Scopus, and Dissertations and Theses (see Appendix B for search syntax). Studies published in any year were considered.

Included studies: (a) were randomized or non-randomized controlled trials (e.g., both designs were included because intervention studies often result in participants being self-selected rather than randomly selected), (b) had student trainee samples (e.g., medical students, dental students, counseling students), (c) were published in English, and (d) were published in scientific, peer reviewed journals or were dissertations or theses. Excluded MI training studies were those that: (a) included only professional trainees (i.e., individuals who have completed their educational training), (b) utilized a single group pre- and post-test design (i.e., no control or comparison group), (c) provided only a narrative review or qualitative analysis, or (d) provided outcome data for only the study clients but not the trainees.

The primary researcher (AM) and a graduate student in Counseling Psychology independently reviewed all of the titles and abstracts from the initial database search, as well as the full-text articles that were selected for further review. The researchers had excellent agreement and substantial reliability ($M_{ICC} = .99$, $M_{Kappa} = 0.80$; see Table 1); the primary researcher and graduate student had adequate reliability and agreement across the full scope of

articles. The two researchers compared included and excluded articles in sections of 1,000 articles at a time. At each set of 1,000 articles, the number of zeros (exclude) and ones (include) were compared, and then differences were discussed. Disagreements regarding inclusion were resolved through discussion between the primary researcher and a graduate student assistant and reviewed by an auditor, as needed. In total, 62 articles were identified for full-text review; four ratings for articles were adjusted between the researchers through discussion and auditing.

Table 1

Cohen's Kappa Agreement for Coders

	Investigator 1 include	Investigator 1 exclude	Total
Investigator 2 include	33	15	48
Investigator 2 exclude	11	9,793	9,804
Total	44	9,808	9,852

Overall Cohen's kappa = .80

Overall rating agreement = 99.9%

$N = 9,852$

Coding

Data were extracted and coded by the primary author and a graduate student in Counseling Psychology. Coded data included demographics, such as the sample's female participant percentage, student degree type (i.e., undergraduate, Master's, doctorate, and student specialty area [e.g., medicine, dentistry, counseling]). Furthermore, primary study characteristics

were extracted. Some of these included: the number of participants in each group (i.e., treatment, control), outcome time points (e.g., baseline, post-training, later follow-up), length of MI training (e.g., two hours), study design (i.e., RCT, nRCT), and type of comparison group (e.g., control, another active treatment). For a complete list of codes, see Appendix C.

Outcomes

Effect sizes were calculated by aggregating outcome measures for available data. Measures used to calculate effect sizes included those collected through direct observation (i.e., rated by independent coders) and student self-report (e.g., MI knowledge). Initially, all outcome effect sizes were combined to determine one overall effect size for MI training effectiveness (i.e., "combined"). Next, effect sizes were separately calculated for eight outcomes commonly assessed in the literature (i.e., empathy, spirit, adherence, open-ended questions, reflections, reflections-to-questions ratio, change plan, knowledge; e.g., Hodorowicz, 2018; Spollen et al., 2010; Young & Hagedorn, 2012). The effect sizes for single outcome behaviors reported in at least six studies (Borenstein et al., 2009; i.e., 8 single behaviors + one combined outcome of the total of the 8 behaviors) were included in the current review. Unexpectedly, most included studies primarily reported independent behavioral outcomes (e.g., number of open-ended questions, percent adherence to MI techniques), rather than total scores on common MI assessment measures (e.g., motivational interviewing treatment integrity; MITI; Moyers et al., 2005). Although the MITI (or an alternative of it: motivational interviewing skill code; MISC) was reported in 10 of the included studies, six studies did not use this measure. Single behaviors were reported in all studies. For studies that had more than one datapoint per outcome measure (e.g., empathy at baseline and post-intervention), each set of data was recorded separately and not assumed to be independent (Cheung, 2019).

MI-Consistent Behavior Measures

MI-consistent behaviors (e.g., open-ended questions, reflections) are commonly measured by independent raters viewing a recorded interview of the trainee with a client (standardized, real, simulated, or through role plays) and tallying specific MI skills use. Researchers often use independent ratings of these behaviors but not an explicit measure (Madson & Campbell, 2006). From the studies included in the analyses, behaviors were reported in three ways; some studies reported global scores (n of outcome datapoints = 35), some reported behavior counts (e.g., number of utterances; n = 13), and some reported a percentage for the amount of time spent engaging in a behavior (n = 6). Each of these forms of data were used in the analyses.

MI Knowledge

In addition to behavioral outcomes, students' knowledge of MI was also commonly assessed (n = 4) among the existing literature. Knowledge about MI was not assessed in all studies, but when it was, either a validated measure (n = 3; e.g., MI Knowledge Assessment Test; MIKAT; Leffingwell, 2006) or several knowledge-based open-ended questions (n = 1; e.g., "What is the spirit of MI?") were utilized. This outcome was exclusively measured through student self-report.

Risk of Bias

Due to the inclusion of both randomized and non-randomized controlled trials, a tool that could be used with both types of study designs was needed. The Evidence Project risk of bias tool (Kennedy et al., 2019; see Appendix D) allows for the assessment of these designs without further adaptation to the tool being needed. The tool was created to assess the quality (i.e., risk of bias) of intervention studies. In total, the tool includes eight items (items 1-3 assesses the study

design, and items 4-8 assesses the quality), rated as "yes" (1) or "no" (0). In some instances, items are rated as "not reported" (NR) or "not applicable" (NA). Items assess factors including: (a) whether the study included a followed cohort, (b) if the study included a control or comparison group, (c) if there was both pre- and post-intervention data, (d) whether there was random assignment of participants to the intervention, (e) whether there was random selection of participants for assessment, (f) if the follow-up rate was 80% or more, (g) if the comparison groups were equivalent on sociodemographics, and (h) if the comparison groups were equivalent at baseline on outcome measures. If a study met the criterion for an item (specific descriptions are available for each item; see Kennedy et al., 2019), it was given a score of 1 ("yes"). If a study did not meet the criterion for an item, it is given a score of 0 ("no"). For the items with ratings of NR and/or NA, items were also given a score of 0. The maximum score a study received was 8 (i.e., 8 items with each earning a "1" or "did include").

In the current study, in order to conduct moderation analyses, it was decided to provide total risk of bias scores for each study (i.e., high risk of bias = 5 or fewer 1s; low risk of bias = 6 or more 1s), as well as a total number of items rated as having high risk of bias for each study (continuous variable). Bias was independently assessed for each included study, and then agreed upon by the primary author and a graduate research assistant. Disagreements were discussed among the two researchers until a consensus was reached.

Statistical Analyses

Comprehensive Meta-Analysis Version 3 (CMA-V3; Borenstein et al., 2009) was used to calculate and analyze effect sizes. Hedges' *g* was used as the effect size measure because it provides an unbiased estimate that is corrected on sample size; it is more precise with a small sample size (Enzman, 2015). Standard effect size interpretations for Hedges' *g* were used: small

(.20 to .49), medium (.50 to .79), and large (.80 or greater; Hedges & Olkin, 1985). Furthermore, a random effects model was used because of the anticipated variability among study effect size distributions. When it is likely that there is not one true effect, a random effects model is preferred (Borenstein et al., 2009).

Data inputted into CMA were descriptive (e.g., means, standard deviations) and/or inferential (e.g., t-, F-, p-values), along with the sample sizes for each group. When the intervention group performed better than the comparison group (e.g., intervention group used more open-ended questions than the comparison), the Hedges' g was positive. If the opposite was true, it was marked as negative. Furthermore, authors of included studies were contacted when essential data (e.g., means, standard deviations, p-values) were not reported. If corresponding authors either did not respond or were not able to provide the necessary data, the study was excluded (this was the case for four studies; see Figure 1).

In order to determine which study outcomes would be included in the final analyses, every possible outcome was extracted from each included study. Multiple conceptually congruent outcomes (e.g., complex and simple reflections; $k = 2$) were averaged, and then combined to allow analyses of additional effect sizes (Scammacca et al., 2014). Studies were then compared to determine how they conceptualized the outcomes to ensure that similar outcome types could be grouped together. If three or more studies reported one outcome (e.g., trainees' MI knowledge), it was included in the analyses, as CMA requires three or more studies in order to conduct meaningful analyses (Borenstein et al., 2009). In sum, eight distinct outcomes were included (see Table 4). As expected, not every study reported each of the eight outcomes. Data for relevant outcomes were extracted from each study; the number of extracted datapoints ranged from one to six. Moreover, assessment timepoints varied from study to study. Many

studies only collected data at post-intervention (e.g., Bray et al., 2013; Daeppen et al., 2012).

Some studies included follow-up outcomes after, but these follow-up effect sizes were not

included because effects likely attenuate with time and there were an insufficient number to

conduct separate analyses. Following Higgins and colleagues (2019), only post-training

measures were used. When pre-training measures were included, they were controlled for their

effects when calculating the post outcome effect size.

Furthermore, potential heterogeneity (e.g., significant dispersion between studies) was

assessed for within the sample, in addition to two a priori (i.e., risk of bias, length of training)

and one exploratory moderator (i.e., type of comparison group). To determine if heterogeneity

was present, a Q statistic (included a p-value) and I^2 were calculated. The Q statistic's p-value

determines its level of significance; however, a significant p-value may be due to a small sample

size. I^2, on the other hand, is not dependent on sample size. Moreover, Bloch (2014) suggests

that I^2 values greater than 20% warrant moderation analyses because this could be indicative of

other contributing factors. Sub-group analysis was used to assess for potential moderators when

the study variable was dichotomous (e.g., type of comparison group). For the continuous

variables (i.e., risk of bias, training length), meta-regression was used. A priori moderators

included quality of study (i.e., risk of bias ratings) and training length (i.e., number of hours).

One other potential moderating variable (i.e., type of comparison group) was exploratory in

nature.

Finally, it is possible that studies were missing from the current review (e.g., unpublished

null findings); this was assessed through publication bias analyses (Borenstein et al., 2009).

Statistics provided included funnel plots and Egger's regression intercept (Duval & Tweedie,

2000; Rothstein, 2005). Both funnel plots and Egger's test were conducted for outcomes with at

least 10 studies (Higgins & Green, 2011). Funnel plots graphically depict study effect sizes, with studies (depicted as closed circles) on both the left and right side of the graph illustrating both positive and null or negative effects. Studies do not always lie on both sides on the graph and may sometimes be outside of the graph's range. In that case, Egger's regression intercept estimates the degree of asymmetry within the data.

Results

Study Selection

A detailed flowchart of the systematic search process can be seen in Figure 1. After excluding duplicate articles, 9,852 articles remained and were reviewed. Titles and abstracts for these records were screened and 9,790 were excluded; a total of 62 records were then reviewed in full-text. Of those, 47 were excluded due to not meeting inclusion criteria (e.g., no control group, unusable data presented and authors unable to be contacted). One duplicate publication was excluded because it used the same sample as another included study but with different analyses. The reported analyses which appeared to be of higher quality was included, following the process suggested by von Elm and colleagues (2004).

Study Characteristics

In total, 15 studies were included in the current meta-analysis and systematic review (see Table 2). This included 1,013 students-in-training; degree types ranged from pursuing a bachelor's degree to doctoral and medical students. The most common participants were medical students ($k = 5$). Other student specialty areas included dental ($k = 3$), counseling ($k = 2$), psychology ($k = 2$), social work ($k = 2$), and pharmacy ($k = 1$). Moreover, outcomes reported in each study varied considerably from one ($k = 1$) to seven outcomes ($k = 2$), as did the number of times outcomes were assessed over time (range = 1-3 assessment points).

Risk of Bias

Table 3 summarizes the outcomes for each study and category of the risk of bias measure. Each of the 15 included studies is plotted at the top of the table, and the eight items on Kennedy and colleagues' (2019) risk of bias scale are listed down the right side of the table. Overall, all included studies met the "Cohort" and "Control or comparison group" items (as indicated by + in the cell). Quality was more mixed for the studies on the remaining scale items. Due to the number of studies that only reported outcome data at post-intervention, fewer studies were able to meet some of the later scale items (as indicated by a NA in the cell). For example, the item "Follow-up rate of 80% or more" could not be assessed in those studies ($k = 6$). Furthermore, it was observed that a number of studies ($k = 10$) failed to report critical pieces of information (e.g., comparison group equivalency, follow-up description). Table 2 also includes a bias categorization (i.e., high, low) for each study. In total, seven studies were categorized as having a "low" risk of bias and the remaining eight studies were deemed to have a "high" risk of bias.

Figure 1

PRISMA Flow Diagram

Identification

Records identified through database searching
(n = 15,083)

Additional records identified through other sources
(n = 429)

Records after duplicates removed
(n = 9,852)

Screening

Records screened
(n = 9,852)

Records excluded
(n = 9,790)

Eligibility

Full-text articles assessed for eligibility
(n = 62)

Full-text articles excluded, with reasons
(n = 47)

No control group (n = 37)

Did not provide usable meta-analysis data (e.g., means, SDs) within manuscript, and study authors either did not respond to requests for data or were not able to provide additional data (n = 4)

Published only data on client outcomes (n = 2)

Dissertation duplicate with published data (n = 1)

Case study (n = 1)

Duplicate with the same sample as an included study (n = 1)

No post-intervention data (n = 1)

Included

Studies included in quantitative synthesis (meta-analysis)
(n = 15)

18

Table 2

Summary of Included Studies

Study	Study population; location	N: Tx/Ctrl	Comparison group	Gender (% female)	Training type; duration (tx/ctrl) hrs.	Time pts.	Type of cts.	Outcomes	RoB [a]
Bray et al. (2013)	Dental hygienist students; U.S.	27/26	Control	NR	Workshops (lecture, practice, coaching); 14/asynchronous	Post-only	Standardized	Empathy	High
Daeppen et al. (2012)	Medical students; Switzerland	42/49	Control	NR	Practical exercises, role plays, videos; 8/0	Post-only	Standardized	Empathy, spirit, MI-adherent bx, open-ended questions, reflections, reflection to question ratio	Low
Haeseler et al. (2011)	Medical students; U.S.	80/19	Control	48%	Lecture role plays, feedback; 2/0	Post-only	Standardized	Empathy, open-ended questions, change plan, reflections	High

19

Study	Population	N	Condition		Training; hrs	Timing	Measure context	Outcomes	Rating
Han & Cho (2018)	Dental hygienist students; South Korea	22/21	Waitlist control	NR	Lecture, role plays, feedback, discussion; 7/0	Post-only	Role plays[b]	Open-ended questions, change plan	High
Hodorowicz (2018)	Social work students; U.S.	9/8	Active treatment	76%	Experiential learning & live supervision OR lecture & coding practice; 12/12	Pre, post, 1 semester f/u	Standardized	Relationship (empathy), reflections, reflection to questions ratio	Low
Klonek & Kauffeld (2016)	Psychology students; Germany	49/28	Control	NR	Lecture, group work, homework, readings; 40/0	Post-only	Recruited	MI-adherent bx, open-ended questions, reflections, reflection to question ratio	High
Koerber et al. (2003)	Dental students; U.S.	11/11	Control	NR	In-class sessions (lecture, practices); 12/0	Pre, post	Standardized	Empathy, change plan	Low

Study	Population	Sample	Condition	Retention	Format/Dose	Design	Assessment	MI Skills Measured	Rating
Madson et al. (2013)	Psychology students; U.S.	Brief = 43, Extended = 25, Intensive = 15	Active treatment	79.5%	In-class lecture & practice (brief to intensive); 40/1	Pre, post	None	Reflective listening (empathy), change plan, knowledge	High
Mounsey et al. (2006)	Medical students; U.S.	47/46	Active treatment	41%	Workshop (didactics, small group practice, demos); 3/3	Post-only	Standardized or role plays [b]	Empathy, spirit, MI-adherent bx, open-ended questions, reflections	High
Opheim et al. (2009)	Medical students; Norway	59/54	Waitlist control	NR	Workshop (practice exercises, student questions, role-plays); 4/0	Post-only	Simulated	Empathy, spirit, open-ended questions, reflections, reflection to question ratio	High

Study	Population	N	Group	%	Intervention	Timepoints	Measure	Outcomes	Risk of bias
Pecukonis et al. (2016)	Social work students; U.S.	26/28	Active treatment	43%	Experiential learning & live supervision OR independent readings & training videos; 14/14	Start of training (T2), post-training (T3), 5-mos f/u (T4)	Standardized	Empathy, spirit, MI-adherent bx, open-ended questions, reflections, reflection to question ratio	Low
Spollen et al. (2010)	Medical students; U.S.	17/18	Control	37%	Workshop (didactics, role-plays); 2/0	Pre, post	Standardized	Empathy, change plan, open-ended questions, knowledge	Low
Widder-Prewett et al. (2017)	Pharmacy students; U.S.	23/21	Active treatment	61%	3 learning sessions & 2 practices times; 6/6	Pre [c], post, & retrospective pre [c]	Standardized	Empathy, spirit, MI adherent bx, open-ended questions	Low

Study	Population; Country	n	Control	%[a]	Training	Assessment	Clients	Outcomes	RoB
Young & Hagedorn (2017)	Counseling students; U.S.	25/18	TAU	88%	Trainings & 3 f/u supervision w/ MI focus OR trainings & f/u supervision w/out MI focus; 7/4	Post-only	Real	Empathy, spirit, MI adherent bx, open-ended questions, reflections, knowledge	Low
Zeligman et al. (2017); U.S.	Counseling students;	29/30	Control	84%	Lecture, role-plays, assessment; 37.5/0	Post-only	None[b]	Reflective listening (empathy), knowledge	High

Note. tx: treatment; ctrl: control; cts.: clients; TAU: treatment-as-usual; LS: live supervision; *f/u*: follow-up; RoB: risk of bias; bx: behavior; NR: not reported

[a] This score was determined based on the total number of 1s each study was given across the eight items. Scores of six or greater were deemed of low risk of bias, and scores of five or fewer were reported to have high risk of bias. [b] With in-training classmates. [c] Data not used in comparisons.

Table 3

Risk of Bias Categorization [a]

Study	Cohort	Control or comparison group	Pre-post intervention data	Random assignment of participants to the intervention	Random selection of participants for assessment	Follow-up rate of 80% or more	Comparison groups equivalent on sociodemographics	Comparison groups equivalent at baseline on outcome measures
Bray et al. (2013)	+	+	+	-	-	NR	NR	NR
Daeppen et al. (2012)	+	+	-	+	+	NA	+	+
Haeseler et al. (2011)	+	+	-	-	+	NA	+	NR
Han & Cho (2018)	+	+	-	+	-	NA	NR	NR
Hodorwicz (2018)	+	+	+	+	+	+	+	+
Klonek & Kauffeld (2016)	+	+	-	-	+	NA	+	NR
Koerber et al. (2003)	+	+	+	-	-	NR	NR	+
Madson et al. (2013)	+	+	+	-	-	NR	+	+
Mounsey et al. (2006)	+	+	-	+	+	NA	-	+
Opheim et al. (2009)	+	+	-	+	NR	NA	+	NR
Pecukonis et al. (2016)	+	+	+	+	+	+	+	+
Spollen et al. (2010)	+	+	+	+	-	+	+	+
Widder-Prewett et al. (2017)	+	+	+	-	NR	+	+	+
Young & Hagedorn (2017)	+	+	+	-	NR	+	+	+
Zeligman et al. (2017)	+	+	+	-	-	NR	+	+

+ criterion met	- criterion not met	NA: not assessed	NR: not reported

[a] In total, 7 studies had "low" risk of bias, and 8 studies had "high" risk of bias.

Outcomes

See Table 4 for effect size and heterogeneity results by study. After combining all outcomes ($k = 15$), analyses indicated a large effect size in relation to the comparison groups, Hedges' $g = 0.90$, 95% CI [0.45, 1.35], $p < .001$ (see Figure 2). Overall, heterogeneity was considerable for all outcomes but MI knowledge.

In terms of the individual outcomes, significant effects were found. The largest was in improving students' knowledge ($k = 4$; $g = 1.26$, 95% CI [0.93, 1.59], $p < .001$; see Figure 3). Furthermore, student empathy was large, $k = 13$; $g = 0.87$, 95% CI [0.38, 1.37], $p = .001$[2] (see Figure 4). MI-trained students also improved in their use of open-ended questions, $k = 10$; $g = 0.67$, 95% CI [0.16, 1.17], $p = .01$ (see Figure 5), and reflections to questions ratio, $k = 5$; $g = 0.70$, 95% CI [0.15, 1.25], $p = .01$.

The remaining outcomes were not found to produce significant results. Students' incorporation of a change plan had a moderate effect ($k = 5$; $g = 0.75$, 95% CI [-0.10, 1.60], $p = .08$). The adherence outcome appeared small in size, $k = 6$; $g = 0.48$, 95% CI [-0.05, 1.00], $p = .07$. The reflection effect size (commonly a combination of any type of MI reflection; e.g., simple, complex) was small, $k = 8$; $g = 0.30$, 95% CI [-0.26, 0.85], $p = .30$.

Table 4

Effect Sizes for Outcomes

Outcome	# studies	Hedges' g	SE	p	95% CI		Q	p	I²
All combined	15	0.90	0.23	< .001	0.45, 1.35		135.20	< .001	89.65
Empathy	13	0.87	0.25	.001	0.38, 1.37		123.69	< .001	90.30
Spirit	6	0.47	0.20	.02	0.07, 0.87		20.81	.001	75.97
Adherence	6	0.48	0.27	.07	-0.05, 1.00		32.76	< .001	84.74
O-E questions	10	0.67	0.26	.01	0.16, 1.17		85.25	< .001	89.44
Reflections	8	0.30	0.28	.30	-0.26, 0.85		67.31	< .001	89.60
Reflections-questions ratio	5	0.70	0.28	.01	0.15, 1.25		22.07	< .001	81.88
Change plan	5	0.75	0.43	.08	-0.10, 1.60		38.80	< .001	89.69
Knowledge	4	1.26	0.17	< .001	0.93, 1.59		3.89	.27	22.83

Note. O-E questions: open-ended questions

26

Figure 2

Forest Plot for Combined Outcomes

Study name	Outcome	Time point	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Bray et al., 2013	Empathy	Post-only	4.427	0.508	0.258	3.431	5.423	8.713	0.000
Deappen et al., 2012	Combined	Post-only	0.925	0.221	0.049	0.492	1.358	4.188	0.000
Haesseler et al., 2011	Combined	Post-only	-0.515	0.257	0.066	-1.018	-0.012	-2.006	0.045
Han & Cho, 2018	Combined	Post-only	1.875	0.361	0.131	1.167	2.584	5.189	0.000
Hodorowicz, 2018	Combined	Pre, post	0.640	0.485	0.236	-0.312	1.591	1.318	0.188
Kbnek & Kauffeld, 2016	Combined	Post-only	1.518	0.267	0.071	0.995	2.041	5.692	0.000
Koerber et al., 2003	Combined	Pre, post	0.230	0.412	0.170	-0.577	1.037	0.559	0.576
Madson et al., 2013	Combined	Pre, post	1.511	0.247	0.061	1.026	1.995	6.111	0.000
Mounsey et al., 2006	Combined	Post-only	-0.066	0.206	0.042	-0.470	0.338	-0.322	0.748
Opheim et al., 2009	Combined	Post-only	0.384	0.198	0.039	-0.004	0.771	1.942	0.052
Pecukonis et al., 2016	Combined	Pre, post	0.419	0.272	0.074	-0.115	0.952	1.537	0.124
Spollen et al., 2010	Combined	Pre, post	0.757	0.346	0.120	0.080	1.435	2.191	0.028
Widor-Prewitt et al., 2017	Combined	Post-only	0.083	0.297	0.088	-0.499	0.665	0.278	0.781
Young & Hagedorn, 2012	Combined	Post-only	0.676	0.319	0.102	0.050	1.301	2.118	0.034
Zeligman et al., 2017	Combined	Post-only	1.284	0.288	0.083	0.719	1.849	4.452	0.000
	Combined	Post-only	0.898	0.229	0.053	0.448	1.348	3.913	0.000

Statistics for each study

Hedges's g and 95% CI

-4.00   -2.00   0.00   2.00   4.00

Favors control   Favors training

Figure 3

Forest Plot for Knowledge Outcome

Study name	Outcome	Time point	Statistics for each study							Hedges's g and 95% CI
			Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value	
Madson et al., 2013	Knowledge	Pre, post	1.490	0.246	0.061	1.007	1.973	6.045	0.000	
Spollen et al., 2010	Knowledge	Pre, post	1.128	0.357	0.127	0.428	1.827	3.159	0.002	
Young & Hagedorn, 2012	Knowledge	Post-only	1.567	0.347	0.121	0.886	2.247	4.511	0.000	
Zeligman et al., 2017	Knowledge	Post-only	0.869	0.270	0.073	0.340	1.398	3.220	0.001	
	Combined		1.258	0.169	0.029	0.926	1.590	7.426	0.000	

-4.00 -2.00 0.00 2.00 4.00

Favors control   Favors training

Figure 4

Forest Plot for Empathy Outcome

Study name	Outcome	Time point	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Bray et al., 2013	Empathy	Post-only	4.427	0.508	0.258	3.431	5.423	8.713	0.000
Deappen et al., 2012	Empathy	Post-only	0.907	0.219	0.048	0.478	1.337	4.141	0.000
Haeseler et al., 2011	Empathy	Post-only	-0.205	0.254	0.064	-0.702	0.292	-0.808	0.419
Koerber et al., 2003	Empathy	Pre, post	0.326	0.413	0.171	-0.484	1.136	0.789	0.430
Mounsey et al., 2006	Empathy	Post-only	0.019	0.206	0.042	-0.384	0.422	0.092	0.927
Opheim et al., 2009	Empathy	Post-only	0.632	0.192	0.037	0.257	1.008	3.298	0.001
Peculkonis et al., 2016	Empathy	Pre, post	0.779	0.279	0.078	0.233	1.325	2.794	0.005
Spollen et al., 2010	Empathy	Pre, post	0.213	0.331	0.110	-0.437	0.863	0.643	0.520
Widder-Prewett et al., 2017	Empathy	Post-only	-0.184	0.297	0.088	-0.766	0.398	-0.619	0.536
Young & Hagedorn, 2012	Empathy	Post-only	1.127	0.327	0.107	0.487	1.768	3.449	0.001
			0.732	0.272	0.074	0.198	1.265	2.687	0.007

Hedges's g and 95% CI

-4.00   -2.00   0.00   2.00   4.00

Favors control   Favors training

Figure 5

Forest Plot for Open-Ended Questions Outcome

Study name	Outcome	Time point	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
					Statistics for each study				
Deappen et al., 2012	Open-ended questions	Post-only	1.391	0.233	0.054	0.935	1.847	5.979	0.000
Haeseler et al., 2011	Open-ended questions	Post-only	-0.588	0.257	0.066	-1.091	-0.085	-2.291	0.022
Han & Cho, 2018	Open-ended questions	Post-only	1.912	0.364	0.132	1.199	2.624	5.259	0.000
Klonek & Kauffeld, 2016	Open-ended questions	Post-only	1.961	0.283	0.080	1.406	2.515	6.933	0.000
Mounsey et al., 2006	Open-ended questions	Post-only	0.091	0.206	0.042	-0.313	0.494	0.440	0.660
Opheim et al., 2009	Open-ended questions	Post-only	0.490	0.190	0.036	0.118	0.862	2.582	0.010
Pecukonis et al., 2016	Open-ended questions	Pre, post	0.197	0.269	0.072	-0.330	0.724	0.733	0.464
Spollen et al., 2010	Open-ended questions	Pre, post	1.125	0.357	0.127	0.426	1.824	3.153	0.002
Widder-Prewett et al., 2017	Open-ended questions	Post-only	0.192	0.297	0.088	-0.391	0.774	0.646	0.519
Young & Hagedorn, 2012	Open-ended questions	Post-only	0.037	0.303	0.092	-0.558	0.632	0.123	0.902
	Open-ended questions	Post-only	0.665	0.256	0.066	0.162	1.168	2.593	0.010

Hedges's g and 95% CI

-4.00 -2.00 0.00 2.00 4.00

Favors control — Favors training

30

Moderators

Heterogeneity appeared to have a significant role in most of the outcomes assessed; therefore, an examination of potential moderators was of particular importance. Of the three potential moderators (risk of bias, MI training length, type of comparison group) examined, only training length was statistically significant (see Table 5).

The length of MI training varied from 2 to 40 hours; therefore, it was hypothesized that training length may have affected the strength of the effect size. A significant moderating effect was found, $Q = 0.03$, $p = .04$, with MI training length explaining at least part of the observed dispersion among studies, as longer training yielded greater effects. On the other hand, no moderating effects were identified with risk of bias ($Q = -0.24$, $p = .10$) or type of comparison. Although differences did exist between groups, none of the between-group analyses were significant.

Publication Bias

To assess for publication bias, funnel plots and Egger's regression intercept were calculated for three outcomes and the combined outcome (at least 10 studies are needed for analyses; Higgins & Green, 2011). The resulting changes to the effect sizes and symmetry were not significant (see Table 6 and Figure 6), suggesting there were likely few missing studies.

Table 5

Moderation via Sub-Group and Meta-Regression Analyses

| | | Sub-group analysis | | | Meta-regression | | | |
Moderator	Category	Hedges' *g*	*Q*-value	*p*-value	Coefficient	SE	95% CI	*p*-value
Risk of bias	---	---	---	---	-0.24	0.14	-0.51, 0.04	.10
Training length	---	---	---	---	0.03	0.01	0.002, 0.06	.04
Comparison group	---	---	1.38	.71	---	---	---	---
	Active treatment	0.51	---	---	---	---	---	---
	Control	1.17	---	---	---	---	---	---
	TAU	0.68	---	---	---	---	---	---
	Waitlist	1.10	---	---	---	---	---	---

Table 6

Publication Bias for Outcomes

Outcome	Adjusted *g*	95% CI	*Q*-value	Eggers intercept	Eggers SE	Eggers CI	Eggers *p* (2-tailed)
Combined	0.60	0.09, 1.10	208.61	5.32	2.91	-0.98, 11.62	.09
Empathy	0.87	0.38, 1.37	123.69	4.22	3.21	-2.86, 11.29	.22
Spirit	0.47	0.07, 0.87	20.81	-0.10	5.21	-14.55, 14.36	.99
O-E questions	0.51	-0.02, 1.04	116.27	4.43	4.83	-6.72, 15.59	.39

Note. O-E questions: open-ended questions

Figure 6

Funnel Plot with Observed Studies for Combined Outcomes

Funnel Plot of Standard Error by Hedges's g
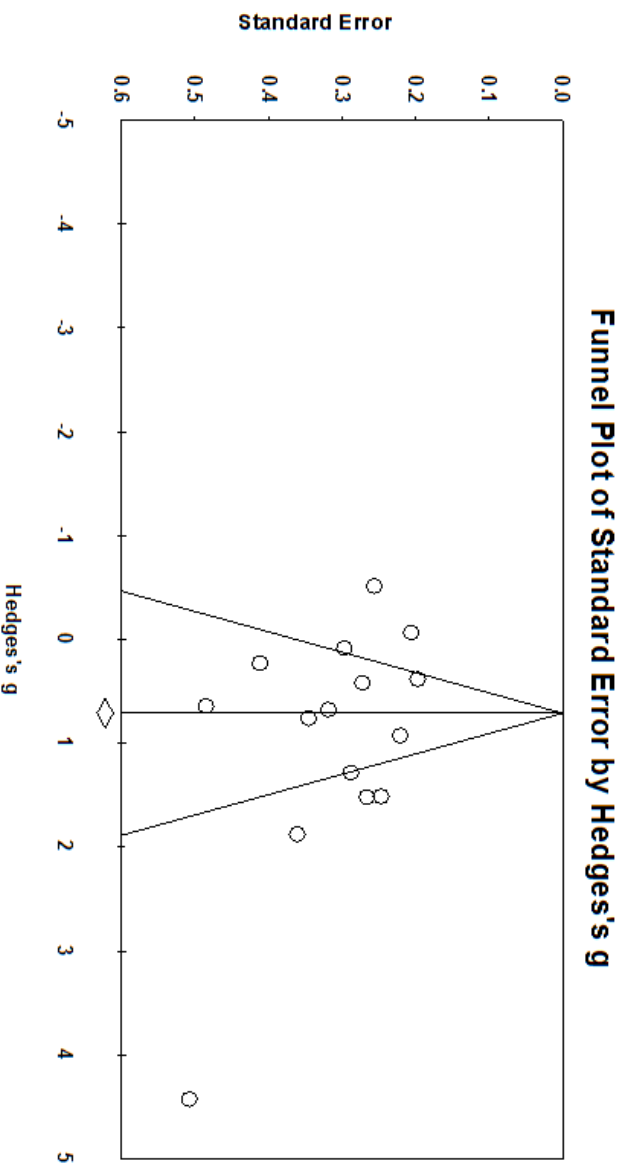
Discussion

The current review examined the effectiveness of teaching students-in-training motivational interviewing. Overall, 15 studies were included in meta-analyses of eight dependent variables yielding a large aggregated effect size (Hedges' $g = 0.90$). A closer examination of the outcomes suggests this effect was likely most affected by moderate-to-large effect sizes for MI knowledge ($g = 1.26$), which is a self-reported outcome, as well as empathy ($g = 0.87$) and open-ended questions ($g = 0.67$), neither of which are MI-specific. Therefore, these large effect sizes should be interpreted cautiously.

This review is the first meta-analysis focusing solely on MI training with a student population. Across outcomes, small-to-moderate (e.g., open-ended questions; $g = 0.67$, reflections to questions ratio; $g = 0.70$, spirit; $g = 0.47$, reflections; $g = 0.30$) to large significant effects (e.g., knowledge; $g = 1.26$, combined; $g = 0.90$, empathy; $g = 0.87$,) were observed. An examination of the effect sizes and plots indicated heterogeneity among all but one outcome. Underpowered analyses may have contributed to some notably sized effects not being statistically significant; thus, caution is warranted when interpreting the specific outcomes.

Given the relatively short duration of some trainings, students may not have received enough time and continued practice to feel confident and comfortable with the more MI-specific skills (e.g., reflections) needed before they could effectively implement the applied skills, which has been identified as a limitation in the literature (Young & Hagedorn, 2012). Although the length of standardized introductory MI trainings tend to vary (one to 24 hours), trainers note that a one-day training is unlikely to result in skills increases and instead, results in familiarity with the material (MINT, 2020). Furthermore, trainers do not suggest that a full three-day MI training will result in reaching skills proficiency; instead, participants tend to have an understanding of

MI and can continue to enhance their skills through ongoing practice (MINT, 2020). In the current study, 12 studies' training length was below this 24 hour threshold.

Moreover, the current results were compared to previous meta-analyses that assessed training professionals in motivational interviewing due to the lack of such reviews with students. The overall effect found in this meta-anlaysis was slightly greater than that found in both Schwalbe and colleagues' (2014) review ($d = 0.76$) and de Roten and colleagues' (2013) study ($g = 0.70$). They did, however, both report high and significant heterogeneity. Surprisingly, previous meta-anlayses have focused more on determining the meaning of the high heterogeneity in their results, rather than determining the effects of other outcomes common to MI (e.g., empathy, reflections, knowledge; Lundahl et al., 2010; Schwalbe et al., 2014). de Roten et al. (2013) noted that professionals' micro-skills (e.g., open-ended questions, reflections) improved the greatest ($g = 0.54$), followed by their empathy ($g = 0.44$). However, the results of the current study suggested that applied micro skills (e.g., reflections, spirit) were more difficult for students to improve than content-based skills (e.g., MI knowledge). One possible reason for this difference is that continued practice is needed for some behaviors (e.g., reflections, reflections to questions ratio) to be able to develop. Given the relatively short duration of training, students may not have received enough time and continued practice to feel confident and comfortable enough with the more basic and foundational skills needed before they could effectively implement the applied skills (Young & Hagedorn, 2012).

Research in the student learning literature support these findings. For instance, this literature supports the need to help students develop lower levels of processing (e.g., knowledge before advancing to higher levels of processing (application; e.g., reflecting). Moreover, empathy

is not a skill that is exclusive to MI; students in many different programs receive training on how to build rapport and be empathetic in other courses and training experiences. Whereas, behaviors such as MI adherence and the reflections to questions ratio may be less prevalent, nonexistent, or inconsistent related to previous training opportunities or experiences (Opheim et al., 2009). For example, medical students may be prone to providing information, rather than reflecting what their patients reported. Additional investigation into these differences is needed to further identify what may be contributing to these results.

Lastly, one of the most significant challenges associated with this study was the heterogeneity observed. It was difficult to determine a pattern or potential reason for this spread. One potential reason for the findings is the inclusion of studies that utilized active treatments for their comparison groups. A closer look at those studies' effects reveals little to no differences between groups and, sometimes, the comparison group displaying better skills than the MI group. Although the inclusion of studies utilizing an active comparison group versus a control or waitlist control adds some differences to aggregation, the basic premise of the current review was to determine how effective MI training was for students, regardless of the type comparison or control group used in the study. Moreover, previous research indicated (Dunhill et al., 2014), and this study supported, the lack of consistency among outcome assessments. Despite there being available measures (e.g., MITI), studies tended to use varying methods of assessing students' skills, attitudes, and knowledge. Taken together, it is likely these factors contributed to the high heterogeneity in the present meta-anlaysis.

Limitations

The current study also includes some limitations. First, the sample size was small, thus

limiting the number of subgroup comparisons that could be analyzed. This also limited the demographic diversity among the sample. Second, heterogeneity among the included studies was very high across outcomes. The lack of significant moderating variables leaves uncertainty about the sources of those differences. It should be noted, however, that it was observed that measurement of the included outcomes varied considerably across studies but primarily included single behavior counts rather than standardized outcome measures. Between the small number of included studies and differences in measurement, it is difficult to further assess this heterogeneity. Third, the 15 included studies assessed skills at different time points; baseline skills (i.e., pre-intervention) were measured in less than half of the included studies (6/15 studies). Therefore, it is difficult to draw conclusions about changes in skills without the knowledge of where students' skills were at prior to receiving the training. Furthermore, few studies included later follow-up investigation beyond post-test ($k = 2$), which prevented an assessment of enduring effects of the training.

One of the goals of training students in MI is to offer them practicical skills to carry over into their future careers (Madson et al., 2013); however, it is uncertain whether students are retaining those skills. Moreover, students at all stages of their educational career – from undergraduates (e.g., Madson et al., 2013) through graduate students (e.g., Young & Hagedorn, 2017) – have received training with no known study including booster or refresher sessions. Pecukonis and colleagues (2016) observed a decay in skills from students who received MI training five-months prior to the follow-up assessment. Hence, it appears that additional investigation into students' long-term skills retention is warranted. Finally, significant variation

was observed in the MI training lengths. Students received as few as two hours of instruction (Haeseler et al., 2011) to as much as 1 semester or 40 hours (Klonek & Kauffeld, 2016; Madson et al., 2013). It was found that training length did affect the amount of information and skills students learned in the training. It would be beneficial to consider a standardized student MI training with, at the least, essential skills that trainings should include – regardless of the students' specialty area.

Future Directions

Despite these limitations, the current review does highlight some important next steps and future directions. Nearly all of the included studies were conducted within a controlled environment (e.g., training clinic), which would allow researchers to assess client impressions of the encounter. However, client impressions were not collected in the included studies; this additional variable would be benefical to include in future studies. It would also be important for student trainees to receive feedback while first in a controlled environment to increase their self-efficacy and enhance their microskills when working with clients. For example, attaining client perceptions of the encounter could be brief, such as a short survey after the session. A corollary to gathering client impressions would be to track client outcomes over time to assess how clients view the trainees' skills development. Notably, this would involve the student meeting with the client on an ongoing basis, rather than just one encounter, as many of the studies used. Tracking client outcomes seems the most feasible with counseling, clinical or counseling psychology, and social work students, as they are the most likely to see clients on a recurring basis. Curiously, however, students in these specialty areas only made up a small fraction of the included studies (6/15 studies). Thus, given that MI appears to be an effective therapeutic technique to encourage behavior change, additional studies involving students in these helping professions appears

warranted. Related, it was interesting to see studies did not target substance use as the primariy change behavior given MI's history in substance use treatment. Although MI has roots in substance abuse treatment (see Miller & Rollnick, 2013; Rollnick et al., 2008), it has since expanded to include many health- or mental health-related changes (Arkowitz et al., 2015; Rollnick et al., 2008). Researchers are encouraged to continue to think of MI as a technique to be used for more than substance use treatment and conduct additional training research related to MI and substance use.

Moreover, a gap in this research appears to be the dearth of studies that included assessment at later follow-up points after the intervention. It is well-established that students tend to quickly forget what they are taught if they do not actively work on retaining and applying it (Murre & Dros, 2015). Thus, methods of encouraging students to actively work on new skills through spaced practice (Dunlosky & Rawson, 2015) and repetition (Dunlosky et al., 2013) may be beneficial for enhancing their long-term retrieval of MI skills. For example, students could identify one or two skills that they struggled with during the training to focus on for a certain duration (e.g., one month), and then check back in with the training instructor to reassess that skill. It would be ideal if instructors followed up and observed students using those skills during this period of time, otherwise students may incorrectly use the skill.

Finally, advances in technology have lent themselves to unique new ways of learning MI skills. One recent pilot study investigated the effectiveness of using telehealth to practice MI skills (Badowski et al., 2019), in which the client encounter occurred via an online synchronous format. In addition, a phone application was created that aids in skills implementation (Vasoya et al., 2019) by tracking certain trainee behaviors (e.g., number of questions, use of reflections) and even alerting trainees when they ask too many closed-ended questions, for example.

Investigation into new developments such as these are needed determine their effectiveness and durability.

Conclusion

Overall, training students in MI appears to be an effective method of teaching students empathy and MI knowledge and helping them develop foundational skills that will likely benefit them in their future professions. The current meta-analysis and systematic review revealed many possible avenues to explore in future research involving student MI training. In particular, future research should include an assessment of student skills at several time points, especially following up after the intervention. At this point, it is unclear if students retain the skills learned in the MI training beyond the intervention period. This literature appears to be a promising area for continued investigation, specifically considering the potential utility of MI skills in most helping-related fields.

Appendix A

PRISMA Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	x
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	1-7
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	8
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	8
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	8-10

42

Section/topic	#	Checklist item	Reported on page #
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	8-10; Appendix B
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	8-10; Appendix B
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	8-10; Appendix B
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	9-11; Table 1; Appendix C
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	11-12
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	12-13; Appendix D
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	13
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I² for each meta-analysis.	13-16
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	15
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	15-16

43

RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	16; Figure 1
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	16; Table 2
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	17; Table 3
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	25; Table 4; Figures 2-5
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	25; Table 4; Figures 2-5
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	31; Table 3
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	31; Table 5; Figure 6
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	35-41
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	37-39
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	39-41
FUNDING			

| Funding | 27 | Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review. | N/A |

45

Appendix B

Search Syntax & Journal Search

Databases: PsycINFO, PsycARTICLES, Consumer Health Complete, PubMed, and Scopus + Dissertations & Theses + journal table of contents sections[a] + Google Scholar

Dates: May, 2019 and September, 2019

Terms: motivational interviewing OR MI AND training OR education AND student* OR trainee* OR intern* OR resident*

Delimited to: peer-reviewed journal articles, theses, and dissertations, published in English

[a]: Patient Education & Counseling, Health Education Journal, Medical Teacher, Counselor Education & Supervision, Medical Education

Appendix C

Coding

- Author name(s)
- Paper title
- Journal name
- Year
- Design
    - o nRCT
    - o RCT
- Location paper was found
    - o Database
    - o Proposal
- Study duration
    - o Variable (2 - 40 hrs.)
- Treatment group total hours
    - o Variable (2 - 40 hrs.)
- Control group total hours
    - o Variable (0 - 14 hrs.)
- Study location
    - o Germany
    - o Norway
    - o South Korea
    - o Switzerland
    - o U.S.
- Study location type
    - o University (all)
- Type of students
    - o Counseling
    - o Dental
    - o Dental hygiene
    - o Medical
    - o Social work
    - o Pediatric resident
    - o Psychology
- Sample size (treatment, control)
- Gender composition
- % female
- Number of groups
    - o 2
    - o 3
- Training group type (most were a combination of these)
    - o Lecture
    - o Exercises and/or homework

- o Experiential
- o Supervision
- o Role plays
- o Workshop
- o Videos
- Time points
  - o Pre, post
  - o Post-only
  - o Pre, after 1st training, after supervision or TAU, 5 mos. follow-up
  - o Post, 1 semester follow-up
  - o 3 and 7 mos. follow-up
- Outcome types
  - o Empathy
  - o Spirit
  - o Adherence
  - o Open-ended questions
  - o Reflections
  - o Reflections-to-questions ratio
  - o Change plan
  - o Knowledge
- Types of clients
  - o None
  - o Recruited, but not standardized
  - o Role plays
  - o Standardized
- Notes

Appendix D

The Evidence Project Risk of Bias Tool (Kennedy et al., 2019)

Risk of bias tool domains	Items	Response choices
Study design	1. Cohort	Yes, no
	2. Control or comparison group	Yes, no
	3. Pre/post intervention data	Yes, no
Participant representativeness	4. Random assignment of participants to the intervention	Yes, no, NA
	5. Random selection of participants for assessment	Yes, no
	6. Follow-up rate of 80% or more	Yes, no, NA, NR
Equivalence of comparison groups	7. Comparison groups equivalent on sociodemographics	Yes, no, NA, NR
	8. Comparison groups equivalent at baseline of outcome measures	Yes, no, NA, NR

References

*References marked with an asterisk indicate studies included in the meta-analysis

Arkowitz, H., Miller, W. R., & Rollnick, S. (2015). *Motivational interviewing in the treatment of psychological problems.* Guilford Press.

Badowski, D. M., Rossler, K. L., & Gill-Gembala, L. (T. (2019). Telehealth simulation with motivational interviewing: Impact on learning and practice. *Journal of Nursing Education*, *58*(4), 221-224. https://doi.org/10.3928/01484834-20190321-06

Barwick, M. A., Bennett, L. M., Johnson, S. N., McGowan, J., & Moore, J. E. (2012). Training health and mental health professionals in motivational interviewing: A systematic review. *Children and Youth Services Review, 34*(9), 1786-1795. https://doi.org/10.1016/j.childyouth.2012.05.012

Bloch, M. (2014). Meta-analysis and moderator analysis: Can the field develop further? *Journal of the American Academy of Child & Adolescent Psychiatry, 53*(2), 135-137. https://doi.org/10.1016/j.jaac.2013.12.001

Bloom, B. S. (1956). Taxonomy of educational objectives: The classification of educational goals. Longmans, Green.

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2009). *Introduction to Meta-Analysis*. John Wiley & Sons. https://doi.org/10.1002/9780470743386

*Bray, K. K., Catley, D., Voelker, M. A., Liston, R., & Williams, K. B. (2013). Motivational interviewing in dental hygiene education: Curriculum modification and evaluation. *Journal of Dental Education, 77*(12), 1662-1669.

Cheung, M. W. L. (2019). A guide to conducting a meta-analysis with non-independent effect

    sizes. *Neuropsychology Review.* Advance online edition. https://doi.org/10.1007/s11065-

    019-09415-6

Cleary, M., Hunt, G., Matheson, S., Siegfried, N., & Walter, G. (2008). Psychosocial

    interventions for people with both severe mental illness and substance misuse. *The*

    Cochrane Database of Systematic Reviews, CD001088.

*Daeppen, J.-B., Fortini, C., Bertholet, N., Bonvin, R., Berney, A., Michaud, P.-A., … Gaume, J.

    (2012). Training medical students to conduct motivational interviewing: A randomized

    controlled trial. *Patient Education and Counseling*, *87*(3).

    https://doi.org/10.1016/j.pec.2011.12.005

de Roten, Y. D., Zimmermann, G., Ortega, D., & Despland, J. (2013). Meta-analysis of the

    effects of MI training on clinicians' behavior. *Journal of Substance Abuse Treatment,*

    45(2), 155-162. https://doi.org/10.1016/j.jsat.2013.02.006

DiClemente, C. C., Corno, C. M., Graydon, M. M., Wiprovnick, A. E., & Knoblach, D. J.

    (2017). Motivational interviewing, enhancement, and brief interventions over the last

    decade: A review of reviews of efficacy and effectiveness. *Psychology of Addictive*

    Behaviors, 31(8), 862-887. https://doi.org/10.1037/adb0000318

Dunhill, D., Schmidt, S., & Klein, R. (2014). Motivational interviewing interventions in graduate

    medical education: A systematic review of the evidence. *Journal of Graduate Medical*

    Education, 6(2), 222-236. https://doi.org/10.4300/jgme-d-13-00124.1

Dunlosky, J., & Rawson, K. A. (2015). Practice tests, spaced practice, and successive relearning: Tips for classroom use and for guiding students' learning. *Scholarship of Teaching and Learning in Psychology, 1*(1), 72-78. https://doi.org/10.1037/stl0000024

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques. *Psychological Science in the Public Interest*, *14*(1), 4-58. https://doi.org/10.1177/1529100612453266

Dunn, C., Deroo, L., & Rivara, F. P. (2001). The use of brief interventions adapted from motivational interviewing across behavioral domains: A systematic review. *Addiction*, *96*(12), 1725-1742. https://doi.org/10.1046/j.1360-0443.2001.961217253.x

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*(2), 455-463. https://doi.org/10.1111/j.0006-341x.2000.00455.x

*Haeseler, F., Fortin, A. H., Pfeiffer, C., Walters, C., & Martino, S. (2011). Assessment of a motivational interviewing curriculum for year 3 medical students using a standardized patient case. *Patient Education and Counseling*, *84*(1), 27-30. https://doi.org/10.1016/j.pec.2010.10.029

Hall, K., Staiger, P. K., Simpson, A., Best, D., & Lubman, D. I. (2016). Continuing the conversation: 30 years of motivational interviewing dissemination. *Addiction, 111*(7), 1158-1159. https://doi.org/10.1111/add.13414

*Han, Y. S., & Cho, Y. S. (2018). Motivational interviewing training using role play in dental hygiene students. *Journal of Dental Hygiene Science,18*(1), 50-59. https://doi.org/10.17135/jdhs.2018.18.1.50

Hedges L. V., & Olkin I. (1985). Statistical methods for meta-analysis. Academic Press.

Higgins, J., & Green, S. (2011). *Cochrane Handbook for Systematic Reviews of Interventions*, Version 5.1.0. The Cochrane Collaboration. Retrieved from www.handbook.chochrane.org

*Hodorowicz, M. T. (2018). Teaching and learning motivational interviewing: Examining the efficacy of two training methods for social work students. Retrieved February 2, 2019, from https://archive.hshsl.umaryland.edu/handle/10713/8007

Hunt, G. E., Siegfried, N., Morley, K., Sitharthan, T., & Cleary, M. (2013). Psychosocial interventions for people with both severe mental illness and substance misuse. *The Cochrane Database of Systematic Reviews, 10*, CD001088.

Iarussi, M. M. (2020). *Integrating motivational interviewing and cognitive behavioral therapy in clinical practice*. Routledge.

Karlsson, P., & Bergmark, A. (2015). Compared with what? An analysis of control-group types in Cochrane and Campbell reviews of psychosocial treatment efficacy with substance use disorders. *Addiction*, *110*(3), 420-428. https://doi.org/10.1111/add.12799

Kennedy, C. E., Fonner, V. A., Armstrong, K. A., Denison, J. A., Yeh, P. T., O'Reilly, K. R., & Sweat, M. D. (2019). The Evidence Project risk of bias tool: Assessing study rigor for both randomized and non-randomized intervention studies. *Systematic Reviews*, *8*(1). https://doi.org/10.1186/s13643-018-0925-0

*Klonek, F. E., & Kauffeld, S. (2016). Watch your language! Analyzing active ingredients of client speech in a motivational interviewing intervention for environmental behavioral change. *Umweltpsychologie*, *20*(1), 62-84.

*Koerber, A., Crawford, J., O'Connell, K. (2003). The effects of teaching dental students brief motivational interviewing for smoking-cessation counseling: A pilot study. *Journal of Dental Education, 67*(4), 439-447.

Leffingwell, T. R. (2006). Motivational interviewing knowledge and attitudes test (MIKAT) for evaluation of training outcomes. *Motivational Interviewing Network of Trainers Bulletin, 13,* 10-11.

Lundahl, B. W., Kunz, C., Brownell, C., Tollefson, D., & Burke, B. L. (2010). A meta-analysis of motivational interviewing: Twenty-five years of empirical studies. *Research on Social Work Practice*, *20*(2), 137-160. https://doi.org/10.1177/1049731509347850

Lundahl, B., Moleni, T., Burke, B. L., Butters, R., Tollefson, D., Butler, C., & Rollnick, S. (2013). Motivational interviewing in medical care settings: A systematic review and meta-analysis of randomized controlled trials. *Patient Education and Counseling, 93*(2), 157-168. https://doi.org/10.1016/j.pec.2013.07.012

Macdonald, P., Hibbs, R., Corfield, F., & Treasure, J. (2012). The use of motivational interviewing in eating disorders: A systematic review. *Psychiatry Research, 200*(1), 1-11

Madson, M. B., & Campbell, T. C. (2006). Measures of fidelity in motivational enhancement: A systematic review. *Journal of Substance Abuse Treatment, 31*(1), 67-73. https://doi.org/10.

Madson, M. B., Loignon, A. C., & Lane, C. (2009). Training in motivational interviewing: A systematic review. *Journal of Substance Abuse Treatment, 36*(1), 101-109. https://doi.org/10.1016/j.sat.2008.05.005.1016/j.jsat.2006.03.010

*Madson, M. B., Schumacher, J. A., Noble, J. J., & Bonnell, M. A. (2013). Teaching

    motivational interviewing to undergraduates. *Teaching of Psychology*, *40*(3), 242-245.

    https://doi.org/10.1177/0098628313487450

Maslowski, A. K. (2019, August 2). A systematic review and meta-analysis of motivational

    interviewing training with students-in-training. Retrieved from osf.io/amvfg

Miller, W. R., & Moyers, T., B. (2006). Eight stages in learning motivational interviewing.

    Journal of Teaching in the Addictions, 5(1). https://doi.org/10.1300/J188v05n01_02

Miller, W. R., & Rollnick, S. (1992). *Motivational interviewing: Preparing people to change*

    addictive behavior. Guilford Press.

Miller, W. R., & Rollnick, S. (2013). *Motivational interviewing: Helping people change*.

    Guilford Press.

Moher D., Liberati A., Tetzlaff J., Altman D. G., & The PRISMA Group (2009) Preferred

    reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS*

    Med, 6(7). https://doi.org/10.1371/journal.pmed.1000097

Motivational Interviewing Network of Trainers. (2020). *Training expectations.* MINT.

    https://motivationalinterviewing.org/training-expectations

*Mounsey, A. L., Bovbjerg, V., White, L., & Gazewood, J. (2006). Do students develop better

    motivational interviewing skills through role-play with standardised patients or with

    student colleagues?. *Medical Education*, *40*(8), 775-780. https://doi.org/10.1111/j.1365-

    2929.2006.02533.x

Moyers, T. B., Manuel, J. K., Wilson, P. G., Hendrickson, S. M. L., Talcott, W., & Durand, P.

> (2007). A randomized trial investigating training in motivational interviewing for
>
> behavioral health providers. *Behavioural and Cognitive Psychotherapy*, *36*(2), 149-162.
>
> https://doi.org/10.1017/s1352465807004055

Moyers, T. B., Martin, T., Manuel, J. K., Hendrickson, S. M. L., & Miller, W. R. (2005).

> Assessing competence in the use of motivational interviewing. *Journal of Substance*
>
> *Abuse Treatment, 28*, 19-26. https://doi.org/10.1016/j.sat.2004.11.001

Murre, J. M. J., & Dros, J. (2015). Replication and analysis of Ebbinghaus' forgetting

> curve. *Plos One*, *10*(7). https://doi.org/10.1371/journal.pone.0120644

*Opheim, A., Andreasson, S., Eklund, A. B., & Prescott, P. (2009). The effects of training

> medical students in motivational interviewing. *Health Education Journal*, *68*(3), 170-
>
> 178. https://doi.org/10.1177/0017896909339454

Palacio, A., Garay, D., Langer, B., Taylor, J., Wood, B. A., & Tamariz, L. (2016). Motivational

> interviewing improves medication adherence: A systematic review and meta-
>
> Analysis. *Journal of General Internal Medicine*, *31*(8), 929-940.
>
> https://doi/.org/10.1007/s11606-016-3685-3

*Pecukonis, E., Greeno, E., Hodorowicz, M., Park, H., Ting, L., Moyers, T., . . . Wirt, C. (2016).

> Teaching motivational interviewing to child welfare social work students using live
>
> supervision and standardized clients: A randomized controlled trial. *Journal of the*
>
> *Society for Social Work and Research, 7*(3), 479-505. https://doi.org/10.1086/688064

Rollnick, S., Miller, W. R., & Butler, C. C. (2008). *Motivational interviewing in health care.*

> Guilford Press.

Rosenman, R., Tennekoon, V., & Hill, L. G. (2011). Measuring bias in self-reported

    data. *International Journal of Behavioural & Healthcare Research*, *2*(4), 320-332.

    https://doi.org/10.1504/IJBHR.2011.043414

Rothstein, H. (2005). *Publication bias in meta-analysis: Prevention, assessment and*

    adjustments. Wiley.

Ryan, R., Hill, S., Prictor, M., & McKenzie, J. (2013). Study quality guide. Retrieved from

    https://cccrg.cochrane.org/sites/cccrg.cochrane.org/files/public/uploads/StudyQualityGui

    de_May%202013.pdf

Scammacca, N., Roberts, G., & Stuebing, K. K. (2014). Meta-analysis with complex research

    designs: Dealing with dependence from multiple measures and multiple group

    comparisons. *Review of Educational Research*, *84*(3), 328-364.

    https://doi.org/10.3102/0034654313500826

Söderlund, L. L., Madson, M. B., Rubak, S., & Nilsen, P. (2011). A systematic review of

    motivational interviewing training for general health care practitioners. *Patient Education*

    and Counseling, 84(1), 16-26. https://doi.org/10.1016/j.pec.2010.06.025

*Spollen, J. J., Thrush, C. R., Mui, D., Woods, M. B., Tariq, S. G., & Hicks, E. (2010). A

    randomized controlled trial of behavior change counseling education for medical

    students. *Medical Teacher,32*(4). https://doi.org/10.3109/01421590903514614

Schwalbe, C. S., Oh, H. Y., & Zweben, A. (2014). Sustaining motivational interviewing: A

    meta-analysis of training studies. *Addiction, 109*(8), 1287-1294.

    https://doi.org/10.1111/add.12558

Vanderwall, F. M. (2015). Impact of motivational interviewing on medication adherence in

    schizophrenia. *Issues in Mental Health Nursing, 36*(11), 900-904.

Vasoya, M. M., Shivakumar, A., Pappu, S., Murphy, C. P., Pei, Y., Bricker, D. A., …
Hershberger, P. J. (2019). ReadMI: An innovative app to support training in motivational interviewing. *Journal of Graduate Medical Education*, *11*(3), 344-346.
https://doi.org/10.4300/jgme-d-18-00839.1

von Elm, E., Poglia, G., Walder, B., & Tramèr, M. R. (2004). Different patterns of duplicate publication: An analysis of articles used in systematic reviews. *JAMA, 291*(8), 974-980.

*Widder-Prewett, R., Draime, J. A., Cameron, G., Anderson, D., Pinkerton, M., & Chen, A. M.
H. (2017). Impact of student vs. faculty facilitators on motivational interviewing student outcomes. *American Journal of Pharmaceutical Education, 81*(6).
https://doi.org/10.5688/ajpe816107

*Young, T. L., & Hagedorn, W. B. (2012). The effect of a brief training in motivational interviewing on trainee skill development. *Counselor Education and Supervision, 51*(2), 82-97. https://doi.org/10.1002/j.1556-6978.2012.00006.x

*Zeligman, M., Dispenza, F., Chang, C. Y., Levy, D. B., McDonald, C. P., & Murphy, T. (2017).
Motivational interviewing training: A pilot study in a masters level counseling program. *Counseling Outcome Research and Evaluation*, *8*(2), 91-104.
https://doi.org/10.1080/21501378.2017.134251