The Technical Hypothesis of Motivational Interviewing: A Meta-Analysis of MI's Key Causal Model

Molly Magill,
Department of Behavioral and Social Sciences, Brown University

Jacques Gaume,
Department of Behavioral and Social Sciences, Brown University

Timothy R. Apodaca,
The Children's Mercy Hospitals and Clinics, University of Missouri, Kansas City

Justin Walthers,
Department of Behavioral and Social Sciences, Brown University

Nadine R. Mastroleo,
Department of Behavioral and Social Sciences, Brown University

Brian Borsari, and
Department of Behavioral and Social Sciences, Brown University, Department of Veteran's Affairs

Richard Longabaugh
Department of Psychiatry and Human Behavior, Brown University

Abstract

Objective—The technical hypothesis of motivational interviewing (MI) posits that therapist implemented MI skills will be related to client speech regarding behavior change and that client speech will predict client outcome. The current meta-analysis is the first aggregate test of this proposed causal model.

Method—A systematic literature review, using stringent inclusion criteria, identified $k = 16$ reports describing 12 primary studies. Review methods calculated the inverse-variance-weighted pooled correlation coefficient for the therapist to client and the client to outcome paths across multiple targeted behaviors (i.e., alcohol or illicit drug use, other addictive behaviors).

Results—Therapist MI-consistent skills were correlated with more client language in favor of behavior change (i.e., change talk; $r = .26$, $p < .0001$), but not less client language against behavior change (i.e., sustain talk; $r = .10$, $p = .09$). MI-inconsistent skills were associated with less change talk ($r = -.17$, $p = .001$) as well as more sustain talk ($r = .07$, $p = .009$). Among these studies,

Correspondence concerning this article should be addressed to: Molly Magill, BoxG-S121-5, Providence RI 02913; molly_magill@brown.edu.
Jacques Gaume is now at Lausanne University Hospital, Lausanne Switzerland.

client change talk was not associated with follow-up outcome ($r = .06$, $p = .41$), but sustain talk was associated with worse outcome ($r = -.24$, $p = .001$). In addition, studies that examined composite client language (e.g., an average of negative and positive statements) showed an overall positive relationship with client behavior change ($r = .12$, $p = .006$; $k = 6$).

Conclusions—This meta-analysis provides an initial test and partial support for a key causal model of MI efficacy. Recommendations for MI practitioners, clinical supervisors, and process researchers are provided.

Keywords

motivational interviewing; change talk; sustain talk; change language; therapy process

Introduction

Motivational interviewing (MI) has been defined as: "a collaborative conversation style for strengthening a person's own motivation and commitment to change" (Miller & Rollnick, 2013, p. 12). MI is grounded in client-centered principles and combines relational (e.g., therapeutic empathy, respect for client autonomy) and technical (e.g., open questions, simple and complex reflections) elements to create a safe, exploratory atmosphere for clients to verbalize personal values, capacities, and reasons regarding behavior change. In MI causal theory, these client statements for or against change (i.e., change and sustain talk, respectively) are hypothesized to mediate intervention efficacy.

While originally designed as a treatment for alcohol use, there have now been several hundred studies published on the use of MI across a variety of behavioral domains. Noonan and Moyers (1997) published the first systematic review of MI in which nine of 11 clinical trials showed efficacy for substance use. In 2001, Dunn, Deroo, and Rivara reviewed 29 studies, expanding earlier work to four target areas: substance use, smoking, HIV risk reduction, and diet and exercise. The strongest support was observed for substance use and the authors concluded that results for other behavioral domains were inconclusive (Dunn et al., 2001). However, a later meta-analysis by Hettema, Steele, and Miller (2005) found a moderate short-term between-group effect size spanning a range of target problems. In a broad summary of the evidence to date, Lundahl and Burke (2009) reviewed four MI meta-analyses, and reported that MI is significantly more effective than no treatment and approximately equal to other viable treatments across nine areas of behavior change. MI has support for efficacy, but effect sizes are generally small to moderate, and vary across delivery contexts, populations, and intervention targets. This indicates a need to better specify how MI is associated with client behavior change.

A causal chain has been defined as a "necessary and sufficient set of variables that will explain the relationship between treatment and treatment outcome" (Longabaugh, 2007 p. 25S). For MI, Arkowitz, Westra, Miller, and Rollnick (2008) proposed three key causal processes: *technical* processes characterized by the impact of therapist skills, *relational* processes characterized by the impact of the overall therapeutic atmosphere and relationship, and *conflict resolution* processes characterized by the impact of successful exploration and resolution of client ambivalence. To date, MI has empirical support for efficacy, and

subsequent studies on these technical (reviewed here), relational, and conflict resolution processes are emerging. For example, relational variables, such as therapist expression of empathy, have shown positive effects on client within-session collaboration and engagement (Moyers, Miller, & Hendrickson, 2005) as well as follow-up alcohol use outcome (Gaume, Gmel, Faouzi, & Daeppen, 2009). Therapeutic alliance with nicotine users can be influenced by the MI-adherence of the therapist (Boardman, Catley, Grobe, Little, & Ahluwalia, 2006). On the other hand, therapist-reported focus on ambivalence has been associated with worse drinking outcome when client readiness to change was low (Magill, Stout, & Apodaca, 2013). Relational and conflict resolution pathways have received less attention than MI technique, yet all of these studies represent a general growth in research on how MI works that has occurred in the past 10 years.

The current meta-analysis is the first aggregate test of the hypothesized technical model of MI efficacy across a range of behavior change targets. Based on the work of Arkowitz and colleagues (2008) and subsequent work by Miller and Rose (2009), we test a causal chain "…that proficient use of the techniques of MI will increase clients' in-session change talk and decrease their sustain talk, which in turn will predict behavior change" (Miller & Rose, 2009, p. 529). In this context, change talk is defined as client language in favor of behavior change (e.g., commitments, reasons, ability to change) and sustain talk is client statements in favor of maintaining the 'status quo' (e.g., commitments, reasons, ability not to change). Additionally relevant to the MI technical model is the proposition that techniques inconsistent with MI, such as confrontations or warnings, will decrease in-session change talk and increase sustain talk, resulting in a lack of behavior change. In this review, we provide aggregate measures of the magnitude and significance of the relationships between therapist MI technical skills and client language (the *a path* of a mediational chain) and between client language and behavior change outcome (the *b path* of a mediational chain). If pooled effect sizes showed heterogeneity between studies, putative study-level moderators were examined. Here, methodological variables expected to systematically relate to effect size were first tested, and in the presence of significant residual heterogeneity, client and treatment factors were considered. The results of this study provide a rigorous test of a central hypothesis of MI causal theory, and offer recommendations for MI practitioners, clinical supervisors, and process researchers.

Method

Study Inclusion

The studies meeting inclusion criteria were English language, peer-reviewed publications. These were studies of Motivational Interviewing (MI), Motivational Enhancement Therapy (MET), and Brief Motivational Interventions (BMI) delivered in a manualized, or other systematic (i.e., session protocol or checklist of suggested discussion areas), format. The MI treatments were delivered to individuals seeking intervention, those opportunistically recruited (e.g., emergency department), or those mandated (e.g., college/university campus alcohol violation) to a behavior change intervention. Finally, inclusion targeted studies that examined *a* and/or *b* path effects using observational coding methods (e.g., Motivational Interviewing Skill Code [MISC; Miller, 2000; Miller Moyers, Ernst, & Amrhein,

2003/2008]; Sequential Code for Observational Process Exchanges [Martin, Moyers, Houck, Christopher, & Miller, 2005]). All behavior change outcomes in the context of an MI intervention were of interest to the present meta-analytic review.

Literature Search

A literature search to obtain all eligible studies was conducted through June of 2012. The first step was a database search in PubMed and PsychInfo with keywords: "change talk", "sustain talk", "client language", "change language", "motivational interviewing skills", "motivational interviewing process", "motivational interviewing mechanisms", "motivational interviewing ingredients". The second step was a hand search of these studies' reference lists, as well as pertinent review papers (i.e., Apodaca & Longabaugh, 2009; Arkowitz et al., 2008; Longabaugh, Magill, Morgenstern & Huebner, 2013; Miller & Rose, 2009) for: (1) additional keywords and (2) any missing studies. The final step was a call for in press papers to: (1) the first authors of all derived studies, (2) identified experts in the area of MI process research, and (3) three listservs relevant to MI, motivational therapies, and process research (Association for Cognitive and Behavioral Therapies; MI International Network of Trainers; Society for Psychotherapy Research). Figure 1 provides a pictorial summary of study inclusion, based on QUORUM guidelines (Moher, Cook, Eastwood, Olkin, Rennie, & Stroup, 1999). Study eligibility was determined by the first author with consensus review by the investigative team. The final meta-analytic sample included 16 reports, describing 12 primary studies, and a corresponding *N* of 1004 individuals.

Coding Methods

Study descriptors—Descriptors of primary study sample characteristics fell into three classes. First, client demographic factors included mean age, percent female participants, primarily Caucasian vs. primarily non-Caucasian sample. Second, client diagnostic factors were treatment seeking vs. non-treatment seeking population, outcome type (alcohol, illicit drug, other behavior, and problem severity (non-dependence or abuse vs. dependence population). The third set of study descriptors, MI implementation factors, were dose (session length in minutes), manualization (treatment manual vs. session protocol), and degree of supervision (report of session recording review, report of group or individual supervision, vs. no report).

Study coding procedure—Each study was assigned an identification number that corresponded to descriptor codes and effect size data within the dataset. Primary study coding was conducted by the first four authors using a combination of independent code and consensus methods. Specifically, study descriptors were independently double coded by the first and fourth author and reliability analyses were conducted (see Table 1). Where descriptor data were missing from primary studies, the original clinical trial upon which the process study was based was reviewed. All descriptor discrepancies were resolved via consensus prior to final data entry. For effect size data at *a* and *b* paths, data extraction decisions, made by the first author, were subjected to consensus review by the second and third authors. All coding guidelines are detailed in a project codebook available from the first author.

Data-analysis

MI technical hypothesis paths—The constructs of interest to this review were summary measures proposed in established MI coding systems (e.g., Houck et al., 2013; Miller et al., 2003/2008) and published within the primary study reports. These variables are typically measured as frequency counts for therapist MI skills (e.g., questions and reflections) and as frequency counts with or without corresponding strength ratings for client change language (e.g., statements of commitment, reason, ability). Specifically, the extracted path effects were between therapist MI micro-skills and client change language (four *a* paths: MI-consistent behaviors to client change talk and to client sustain talk; MI-inconsistent behaviors to client change talk and to client sustain talk) and between client change language and client behavioral outcome (three *b* paths: client change talk to outcome, client sustain talk to outcome, and composite of client change language [i.e., combined frequency or averaged strength of negative and positive statements]). Outcomes were extracted based on the primary targeted behavior change within the report and at the earliest reported time point, since few studies provided data for multiple follow-up time points (Baer, Beadnell, Hartzler, Wells, & Peterson, 2008; Hodgins, Ching, & McEwen, 2009). When studies reported negative outcomes (e.g., number of days drinking), these effects were reverse scored such that all estimates indicated a prediction of positive behavior change. Because effect size data were derived from observationally-rated summary scores, only those that achieved "fair" or higher (as defined by Landis & Koch, 1977) inter-rater reliability within the original study were extracted for meta-analytic review.

Path effect size extraction, preparation, and analysis—The effect size for the current study was the pooled correlation coefficient, which provides an inverse-variance-weighted indicator of the significance, strength, and direction of the bivariate relationship of interest. Where needed, primary study statistics (e.g., *t, F, OR*) were transformed into *r* (see e.g., Lipsey and Wilson, 2001 for formulae), and all effect estimates were *z*-transformed for analyses and returned to the *r* metric for reporting purposes. Eight study authors were contacted and asked to provide specific correlational data, and six responded with this information.

The MI causal paths were considered *random effects* from a distribution of studies with both known and unknown moderators of effect magnitude. The *random effects* method re-weights individual effect sizes by adding a constant that represents population variability, providing a more conservative estimate of significance and allowing broader generalization to the population of studies from which the effect sizes were derived (Hedges & Vivea, 2001). Sensitivity analyses were conducted that re-pool effect size data with 'one-study-removed', and trimmed estimates without influential studies (i.e., those that if removed, would change the substantive conclusion of the review) were presented along with each pooled effect size (Baujat, Mahe, Pignon, & Hill, 2002). In addition, the *Q* statistic tested for the presence of significant between-study heterogeneity, and when the *Q* value was statistically significant, study-level moderators were examined (Higgins & Thompson, 2002). Specifically, Maximum Likelihood regression models (using METAREG; Wilson, 2005) first tested methodological control variables potentially predictive of the *a* path (i.e., a temporally ordered sequential test vs. correlational test, length of session in minutes) or the *b* path (i.e.,

a covariate adjusted test vs. not a covariate adjusted test, time point for follow-up assessment) effect size. Next, if there was statistically significant residual heterogeneity, substantive client- and treatment-level moderator variables were tested (see Study Descriptors above). Overall, the aim was to derive homogeneous path effect sizes therefore increasing confidence that the population of studies testing the relationships of interest had been specified. With the exception of reliability analyses (SPSS 20.0) and multivariate meta-regression analyses (SAS 9.2), this review was conducted in Comprehensive Meta-Analysis Version 2.0.

Results

Sample of Primary Studies

A total of 16 reports described 12 primary studies examining the *a, b,* or *a* and *b* paths of the MI technical hypothesis. Study sample descriptor data are provided in Table 1. These studies included a total of 1004 individuals that were primarily adult alcohol users. Eight of 12 studies reported sample race/ethnicity data, and the majority of these were primarily Caucasian samples, including two Swiss samples. The mean age across samples was 35 (*SD* = 10) and the mean proportion of female participants was .40 (*SD* = .18; range 0 to .60). There was a fairly even distribution for treatment compared to non-treatment seeking populations and the majority of studies targeted 'problematic' use behavior rather than meeting some criterion for dependence. The MI, BMI, or MET sessions ranged from 16 to 80 minutes in length (*M* = 47[*SD* = 20] minutes). Two of 12 studies examined MET (Campbell, Adamson, & Carter, 2010; Moyers Martin, Christopher, & Tonigan, 2009) and therefore involved four therapy sessions. However, only one of these studies coded all four MET sessions (Campbell et al., 2010) and these effects were averaged over the course of treatment in the present review. The MI interventions were equally distributed with regard to delivery guided by a session protocol or topic checklist versus use of a treatment manual. Of the studies that described a supervision protocol (*k* = 9), all reported some form of therapy session review (i.e., session audio recording or transcript). Finally, all descriptor data were double coded reliably, as indicated by Intraclass Correlation Coefficients for continuous variables and Kappa Coefficients for categorical variables, with the exception of the client use severity variable. This indicator showed "fair" agreement (Landis & Koch, 1977), and was not considered for moderator analyses (see Table 1).

Tables 2, 3, and 4 contain primary study effect sizes, along with a selection of key characteristics (i.e., sample size, sequential *a* path, target behavioral outcome, covariate adjustment at *b* path, time of outcome follow-up).

Therapist MI Skills in Relation to Client Change Language – "a" path

Change talk—The technical hypothesis of MI proposes that greater therapist use of MI-consistent skills (e.g., open questions, simple and complex reflections, affirmations) will result in greater change talk. In contrast, when therapists exhibit behaviors inconsistent with MI (e.g., confrontations, warnings, unsolicited advice), change talk is proposed to be reduced. The positive MI-consistent to change talk path was supported across seven primary studies that contributed eight effect sizes (Vader, Walters, Prabhu, Houck, & Field, 2010

examined two different MI conditions). Specifically, the inverse-variance-weighted pooled correlation coefficient was $r = .26$ (95% CI [.16, .35]; $p < .0001$, $Q < .05$). Moreover, MI-inconsistent therapist behaviors were negatively associated with client statements in favor of change ($r = -.17$, 95% CI [−.26, −.07]; $p = .001$, $k = 6$, $Q < .05$). In these results, sensitivity analyses found no influential studies. Therefore, the hypothesized relationships between therapist skills and client change talk were supported in our sample of studies, but these effects contained residual between-study heterogeneity.

Sustain talk—The MI technical hypothesis proposes that MI-consistent skills should result in lower sustain talk. Here and contrary to that proposed, the random effects pooled estimate was positive, non-significant, and heterogeneous ($r = .10$, 95% CI [−.02, .22]; $p = .09$, $k = 8$, $Q < .05$). In addition, this positive effect became statistically significant ($r = .14$, $p = .004$) in sensitivity analyses with one, high weight, negative effect study (Moyers, Martin, Christopher, & Tonigan, 2009) removed. Finally, when the path from MI-inconsistent skills to greater sustain talk was examined, the random effects pooled correlation coefficient was positive, significant, and homogeneous ($r = .07$, 95% CI [.02, .13]; $p = .009$, $k = 6$, $Q > .05$). Thus, the *a* path data support the hypothesized links of MI-inconsistent skills to client change and sustain talk, while only the link from MI-consistent skills to change talk was clearly demonstrated.

Client Change Language in Relation to Client Follow-up Outcomes – "b" path

Change and sustain talk—The technical hypothesis of MI proposes that client statements for (change talk) and against (sustain talk) changing the targeted behavior are key causal mechanisms of MI. Across behavior change outcomes (i.e., continuously measured indicators of alcohol use, illicit drug use, or other addictive behaviors), the random effects pooled correlation for change talk was positive, non-significant, and heterogeneous ($r = .06$, 95% CI [− .09, .21]; $p = .406$, $k = 7$, $Q < .05$) while the sustain talk path was negative, significant, and homogeneous ($r = -.24$, 95% CI [− .36, −.11]; $p = .001$, $k = 9$, $Q > .05$). In sensitivity analyses, the *b* path effect for change talk approached statistical significance ($r = .11$, $p = .045$) with one high weight, negative effect study (Vader et al., 2010; MI-without feedback condition) removed. Therefore, client change talk was not a conclusive predictor of reductions in the problematic behavior at follow-up, but client sustain talk predicted worse outcomes.

Composite change language—In studies that examine the relationship between client language and outcome, the construct of client change language has also been measured as a composite indicator along a negative to positive continuum. Typically, these studies have examined total change language strength, commitment strength only, or combined frequency of positive and negative statements. Such composite indicators can be interpreted as a single measure of motivational balance rather than two sides of the ambivalence, each with a putative independent effect. Here, the composite change language to outcome path was positive, significant, and homogeneous ($r = .12$, 95% CI [.03, .21]; $p = .006$, $k = 6$, $Q > .05$). Moreover, no influential studies were observed in sensitivity analyses.

Figure 2 provides a summary view of support for the MI technical hypothesis. The pooled effect sizes supported five of seven hypothesized causal paths. Specifically, therapist MI-consistent skills were associated with more client change talk, but not less sustain talk. In addition, therapist MI-inconsistent skills were associated with less client change talk as well as more sustain talk. Regarding client language as a mechanism of MI efficacy, change talk was not significant, but sustain talk was associated with poor outcome. Finally, when examined as a composite measure of both negative and positive statements, the overall construct of client change language was predictive of positive behavior change.

Moderators of Between-Study Variability in MI Path Effect Sizes

In this meta-analytic review, four of the seven pooled effect sizes showed between-study heterogeneity as indicated by significant Q test values. A random effects model will provide a conservative measure of statistical significance, but the presence of heterogeneity suggests there may be systematic variation in effect size magnitude that has not been fully specified. Here, moderator results for therapist MI-consistent skills to client language showed that effect size variability was explained by smaller observed effects in analyses of sequential associations compared to correlational associations for change ($b = -.49, p < .001$; sequential $r = .09, p = .01$ vs. correlational $r = .52, p = .001$) and sustain ($b = -.46, p = .002$; sequential $r = -.01, p = .10$ vs. correlational $r = .24, p = .08$) talk, and for sustain talk, session length in minutes was also associated with smaller a path effects ($b = -.01, p = .006$). Because these meta-regression models resulted in non-significant residual heterogeneity, no client- or treatment-level variables were examined. For the MI-inconsistent to change talk path, sequential compared to correlational effects and session length in minutes, as well as our client-and treatment-level variables did not predict between-study variability. The effect of client change talk on outcome also did not show any moderator effects. Overall, these results suggest systematically smaller effect sizes for MI-consistent skills to change language paths in studies that incorporate temporally lagged, sequential rather than session-level, correlational analyses.

Discussion

This meta-analysis supports five of the seven proposed paths of the technical hypothesis of MI efficacy. Therapist MI-consistent skills were associated with higher rates of client change talk, while therapist MI-inconsistent skills were associated with less change talk and more sustain talk. Higher rates of within-session sustain talk were associated with worse client outcome. Regarding the key proposed mechanism of MI, change talk, only the effect size for studies that tested a combined measure of change and sustain talk showed an overall positive relationship with behavior change, while the independent effect for change talk was non-significant. We now consider these findings in further detail.

MI prescribed, and therapist enacted, technical skills such as the use of reflective listening techniques, exploratory open questions, and affirming statements were positively associated with client statements in favor of behavior change. While effect sizes were smaller in temporally ordered compared to correlational associations, they remained significant, providing sound support for these behaviors' effect on MI's purported key mechanism. MI-

prescribed skills, however, did not reduce client sustain talk and only three of eight primary studies supported the proposed negative relationship between these variables (Apodaca, Magill, Longabaugh, Jackson, & Monti, 2013; Miller, Benefield, & Tonigan, 1993; Moyers et al., 2009). This finding may speak to a limitation of the current meta-analytic method for examining what is likely, a dynamic relationship. Specifically, MI has been defined as a supportive yet directive counseling approach aimed at exploring and resolving client ambivalence regarding behavior change (Miller & Rollnick, 2002). Therapist MI-consistent skills may therefore increase *both* change and sustain talk, and particularly in contexts where clients are ambivalent. An average of correlations or sequential probabilities may not capture a clinical process that can first *heighten*, but then *resolve*, ambivalence regarding change. Future research should carefully consider these elements of MI process by examining how therapist interventions impact different trajectories of client speech.

Although typically observed in low frequency relative to MI-consistent behaviors, MI-inconsistent behaviors such as confrontations, warnings, or unsolicited advising appeared to be particularly harmful to a motivational interview. This was demonstrated through their relationship to client sustain talk, and the subsequent relationship between sustain talk and poor client outcome. In their 2009 review of self-reported and observer-rated mechanisms of MI, Apodaca and Longabaugh also found support for the detrimental effect of MI-inconsistent therapist behaviors and only one study from the present report was included in this prior review (Miller et al., 1993). Originally, MI-inconsistent behaviors were delineated to highlight the distinction between MI practices and traditional chemical dependency counseling (Miller, 2000). Therefore, in clinical trials of supervised and manualized MI, it is not surprising that MI-inconsistency was rare. The results of the present review underscore the importance of attending to MI-inconsistent behaviors in MI training as they can have a detrimental impact even when occurring only a few times within a session.

Perhaps the most unexpected finding was that change talk, when assessed as an independent frequency measure, did not significantly affect client outcome. In this review, change talk was significantly and positively related to outcome in two of seven studies (Hodgins et al., 2009; Moyers et al., 2009) while the effect size was significant and negative in one study with college students receiving MI without personalized feedback (Vader et al. 2010). With this outlier study removed, there was a small positive effect for change talk that was significant. Vader and colleagues (2010) suggest that personalized educational feedback is especially important for resolving client ambivalence in a young adult, non-treatment seeking population. Further, in two studies that also worked with young, non-treatment seekers, specific sub-dimensions of change talk rather than the summary measure were supported in the published reports (i.e., reason to change among homeless adolescents, Baer et al., 2008, and desire/ability/need to change among young adult, Swiss, males, Gaume, Bertholet, Faouzi, Gmel, & Daeppen, 2013). Therefore, the validity of the MI technical model may depend on particular contexts such as client age or treatment seeking status.

Rather than in independent frequency measure, our results suggest that change talk provides an accurate indicator of intention when considered in the presence of sustain talk. MI scholars have proposed varying measurement approaches to testing the MI technical model, but a methodological standard has not been established. For example, commitment has also

been framed as the key casual predictor and other statements, such as reasons or desires, are only those that should mobilize commitment (Amrhein Miller, Yahne, Palmer, & Fulcher, 2003; Amrhein, 2004; Miller & Rose, 2009; Miller & Rollnick, 2013). In the present review, four of six studies reported averaged commitment strength as the primary change language measure (Amrhein et al., 2003; Campbell et al., 2010; Hodgins et al., 2009; Morgenstern Kuerbis, Amrhein, Hail, Lynch, & McKay, 2013). While this meta-analysis supports a single construct of change language along a negative to positive continuum over two independent effects, further research is also needed to identify the optimal measurement approach for testing client decision-making as a mechanism of behavior change in the context of MI.

At present, our findings indicate clinicians should not focus exclusively on the presence and frequency of change talk, but rather consider the mechanistic role of client speech as a balance of pro- and anti-change statements. The question is how. The current aggregate representation of MI mechanisms could not capture how clinicians are reacting or should react to client ambivalence, hesitance, and/or resistance, as measured here in the form of sustain talk. MI-inconsistent behaviors, though rare, produce more sustain talk, while MI-consistent behaviors both increase and decrease it, depending on the dynamic of the session. Therefore, a preliminary causal chain from MI-inconsistent behaviors to sustain talk and from sustain talk to poor outcomes is supported by the current data, and this is consistent with previous review (Apodaca & Longabaugh, 2009). Yet, this describes what not to do, rather than what to do. In conjunction with our results on composite change language, the central question that arises is whether sustain talk should be avoided due to its negative effect, or if the literature at present has not sufficiently addressed the distinction between explored and resolved ambivalence.

Limitations

The present review has some limitations to consider. These limitations are regarding the measurement of key predictors, the comprehensiveness of the theoretical model tested, and the size of the primary study sample. First, with respect to measures used to examine a and b path effects, it is important to note that each composite variable (therapist MI-consistent and MI-inconsistent skills and client change talk and sustain talk) was comprised of numerous sub-codes identified in the original MI coding systems (Miller, 2000; Miller et al., 2003/2008). Sub-codes of observer-rated MI variables are typically collapsed into broader categories in the process literature, but the relative predictive validity of intervention or language sub-types could not be tested in the present review. For example, a therapist's simple reflection (e.g., "You're not happy about this") of a client statement would carry the same weight as a complex reflection (e.g., "You're at the end of your rope with all of this, and you think you're ready to finally get serious about making changes") of the same client statement. Yet, the effect of these behaviors on the client could easily differ. Likewise, sub-dimensions of change talk may have varying importance for client sub-populations or clinical contexts. We tested these possible moderator effects, but the findings may have been hindered by the size of our sample. We consider attention to the variable processes that may be captured within the MI coding systems (i.e., MISC) as a vital direction for future MI process research.

A second limitation is the focus on MI technical factors without taking into account the broader *relational context* described by Miller and Rose (2009) or the broader *conflict resolution context* described by Arkowitz and colleagues (2008). Analyses using relational predictors such as session-level therapist interpersonal skills as putative conditions moderating MI technical skills' impact on client decision-making is an additional direction worth pursuing. The present findings on the person-level composite change language variable also suggest a need for more nuanced attention to MI's key mechanism of change. Is enhanced motivation (as measured by increased client change talk) or resolved ambivalence (as measured by an increasingly positive balance of pro- and anti-change statements) the optimal indicator of what carries the benefit of a 'motivational interview'? Finally, there was residual heterogeneity in some of the paths analyzed in the present review. Within this relatively large client sample (N=1,004), populations, contexts, and methods differed across studies. Meta-regression techniques, considering study-level moderators, explained the heterogeneity observed in two of four paths with significant Q values. Therefore, future meta-analytic review when a larger number of studies are available will be important to confirm and expand upon the results presented here. Such a review will have an even broader impact if more process analyses across behavior domains are undertaken.

Conclusions

This study offers an aggregate test and critical evaluation of an important emerging area of research: the proposed technical hypothesis of MI efficacy. Noteworthy is that five of the seven predictive paths hypothesized by MI theorists were supported, providing preliminary evidence that MI is operating as proposed. While supported, however, these path effect sizes were small. We believe this is due to a need to enhance the sensitivity of MI process research in some of the ways noted here, but the present results additionally suggest there are key MI mechanisms going unmeasured and untested. Attending to these considerations should offer fruitful results in this next phase of the MI process research literature.

Acknowledgments

References

Amrhein PC. How does motivational interviewing work? What client talk reveals. Journal of Cognitive Psychotherapy. 2004; 18(4):323–336.10.1891/jcop.18.4.323.64001

*. Amrhein PC, Miller WR, Yahne C, Palmer M, Fulcher L. Client commitment language during motivational interviewing predicts drug use outcomes. Journal of Consulting and Clinical Psychology. 2003; 71(5):862–878. [PubMed: 14516235]

Apodaca T, Longabaugh R. Mechanisms of change in motivational interviewing: A review and preliminary evaluation of the evidence. Addiction. 2009; 104(5):705–715. [PubMed: 19413785]

*. Apodaca TR, Magill M, Longabaugh R, Jackson KM, Monti PM. Effect of a significant other on client change talk in motivational interviewing. Journal of Consulting and Clinical Psychology. 2013; 81(1):35–46.10.1037/a0030881. [PubMed: 23231575]

Arkowitz, H.; Miller, WR.; Westra, HA.; Rollnick, S. Motivational Interviewing in the treatment of psychological problems: Conclusions and future directions. In: Arkowitz, H.; Westra, HA.; Miller, WR.; Rollnick, S., editors. Motivational Interviewing in the Treatment of Psychological Problems. Vol. Ch 13. Guilford; New York: 2008. p. 324-41.

*. Baer JS, Beadnell B, Hartzler B, Wells EA, Peterson PL. Adolescent change language within a brief motivational intervention and substance use outcomes. Psychology of Addictive Behaviors. 2008; 22:570–575. [PubMed: 19071983]

Baujat B, Mahe C, Pignon JP, Hill C. A graphical method for exploring heterogeneity in meta-analysis: Application to a meta-analysis of 65 trials. Statistics in Medicine. 2002; 21(18):2641–52. [PubMed: 12228882]

*. Bertholet N, Faouzi M, Gmel G, Deappen JB. Change talk sequence during brief motivational intervention, towards or away from drinking. Addiction. 2010; 105:2106–2112. [PubMed: 20840175]

Boardman T, Catley D, Grobe JE, Little TD, Ahluwalia JS. Using motivational interviewing with smokers: Do therapist behaviors relate to engagement and therapeutic alliance? Journal of Substance Abuse Treatment. 2006; 31(4):329–339.10.1016/j.jsat.2006.05.006 [PubMed: 17084786]

*. Campbell SD, Adamson SJ, Carter JD. Client language during motivational enhancement therapy and alcohol use outcome. Behavioral and Cognitive Psychotherapy. 2010; 38:399–415.

*. Catley D, Harris K, Mayo MS, Hall S, Okuyemi KS, Boardman T, Ahluwalia JS. Adherence to principles of motivational interviewing and client within-session behavior. Behavioral and Cognitive Psychotherapy. 2006; 34(1):43–56.10.1017/S1352465805002432

Dunn C, Deroo L, Rivara FP. The use of brief interventions adapted from motivational interviewing across behavioral domains: A systematic review. Addiction. 2001; 96(12):1725–1742.10.1080/09652140120089481 [PubMed: 11784466]

*. Gaume J, Bertholet N, Faouzi M, Gmel G, Daeppen JB. Counselor motivational interviewing skills and young adult change talk articulation during brief motivational interventions. Journal of Substance Abuse Treatment. 2010; 39:272–281. [PubMed: 20708900]

*. Gaume J, Gmel G, Faouzi M, Daeppen J. Counsellor behaviours and patient language during brief motivational interventions: A sequential analysis of speech. Addiction. 2008a; 103(11):1793–1800.10.1111/j.1360-0443.2008.02337.x [PubMed: 19032529]

*. Gaume J, Gmel G, Daeppen J. Brief Alcohol Interventions: Do Counsellors' and Patients' Communication Characteristics Predict Change? Alcohol & Alcoholism. 2008b; 43(1):62–69.10.1093/alcalc/agm141 [PubMed: 17942439]

*. Gaume J, Bertholet N, Faouzi M, Gmel G, Daeppen JB. Does Change Talk During Brief Motivational Interventions with Young Men Predict Change in Alcohol Use? Journal of Substance Abuse Treatment. 2013; 44(2):177–185. [PubMed: 22658289]

Gaume J, Gmel G, Faouzi M, Daeppen J. Counselor skill influences outcomes of brief motivational interventions. Journal of Substance Abuse Treatment. 2009; 37(2):151–159.10.1016/j.jsat.2008.12.001 [PubMed: 19339147]

Hedges LV, Vevea JL. Fixed- and random-effects models in meta-analysis. Psychological Methods. 1998; 3(4):486–504.10.1037/1082-989X.3.4.486

Hettema J, Steele J, Miller WR. Motivational Interviewing. Annual Review of Clinical Psychology. 2005; 1(1):91–111.

Higgins JP, Thompson SG. Quantifying heterogeneity in meta-analysis. Statistics in Medicine. 2002; 21:1539–1558. [PubMed: 12111919]

*. Hodgins DC, Ching LE, McEwen J. Strength of commitment language in motivational interviewing and gambling outcomes. Psychology of Addictive Behaviors. 2009; 23(1):122–130. [PubMed: 19290696]

Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977; 33:159–174. [PubMed: 843571]

Lipsey, MW.; Wilson, DB. Practical Meta-Analysis. Thousand Oaks, CA US: Sage Publications, Inc; 2001.

Longabaugh R. The search for mechanisms of change in behavioral treatments for alcohol use disorders: A commentary. Alcoholism: Clinical and Experimental Research. 2007; 31(S3):21S–32S.

Longabaugh, R.; Magill, M.; Morgenstern, J.; Huebner, R. Mechanisms of behavior change in treatment for alcohol and other drug use disorders. In: McCrady, BS.; Epstein, EE., editors. Addictions: A Comprehensive Guidebook. 2. Vol. Ch 27. USA: Oxford University Press; 2013.

Lundahl B, Burke BL. The effectiveness and applicability of motivational interviewing: A practice-friendly review of four meta-analyses. Journal of Clinical Psychology. 2009; 65(11):1232–1245.10.1002/jclp.20638 [PubMed: 19739205]

Magill M, Stout RL, Apodaca TR. Therapist effects on client drinking across four motivational interviewing sessions: A longitudinal analysis of process predictors. Psychology of Addictive Behaviors. 2013; 27(3):54–62.

Martin, T.; Moyers, TB.; Houck, J.; Christopher, P.; Miller, WR. Unpublished Manual. 2005. Motivational Interviewing Sequential Code for Observing Process Exchanges (MI-SCOPE) Coder's Manual.

Miller, WR. Unpublished Manual. 2000. Manual for the Motivational Interviewing Skill Code (MISC).

*. Miller WR, Benefield GR, Tonigan JS. Enhancing motivation for change in problem drinking: A controlled comparison of two therapist styles. Journal of Consulting and Clinical Psychology. 1993; 61(3):455–461. [PubMed: 8326047]

Miller, WR.; Moyers, TB.; Ernst, D.; Amrhein, P. Unpublished Manual. 2003. Manual for the Motivational Interviewing Skill Code Version 2.0 (MISC).

Miller WR, Moyers TB, Ernst D, Amrhein P. Manual for the Motivational Interviewing Skill Code Version 2.1 (MISC). Unpublished Manual. 2008

Miller, WR.; Moyers, TB.; Manuel, JK.; Christopher, P.; Amrhein, P. Unpublished Manual. 2008. Revision for Client Language Coding: MISC 2.1: Client Language Assessment in Motivational Interviewing (CLAMI) Segment.

Miller, WR.; Rollnick, S. Motivational Interviewing: Preparing People for Change. 2. New York: Guilford; 2002.

Miller, WR.; Rollnick, S. Motivational Interviewing: Helping People Change. 3. New York, NY: Guilford Press; 2013.

Miller WR, Rose GS. Toward a theory of motivational interviewing. American Psychologist. 2009; 64:527–37. [PubMed: 19739882]

Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomized controlled trials: he QUOROM statement. Quality of Reporting of Meta-analyses. Lancet. 1999; 354:1896–1900. [PubMed: 10584742]

*. Morgenstern J, Kuerbis A, Amrhein P, Hail L, Lynch K, McKay JR. Motivational interviewing: A pilot test of active ingredients and mechanisms of change. Psychology of Addictive Behaviors. 2012; 26(4):859–69. [PubMed: 22905896]

*. Moyers TB, Martin TB, Christopher PJ, Houck JM, Tonigan JS, Amrhein PC. Client language as a mediator of motivational interviewing efficacy: Where is the evidence? Alcoholism: Clinical and Experimental Research. 2007; 31(3):40S–47S.

*. Moyers TB, Martin T, Christopher PJ, Tonigan JS. From in-session behaviors to drinking outcomes: A causal chain for motivational interviewing. Journal of Consulting and Clinical Psychology. 2009; 77(6):1113–1124. [PubMed: 19968387]

Moyers TB, Miller WR, Hendrickson SL. How does motivational interviewing work? Therapist interpersonal skill predicts client involvement within motivational interviewing sessions. Journal of Consulting and Clinical Psychology. 2005; 73(4):590–598.10.1037/0022-006X.73.4.590 [PubMed: 16173846]

Noonan WC, Moyers TB. Motivational interviewing: A review. Journal of Substance Misuse. 1997; 2:8–16.

*. Vader AM, Walters ST, Prabhu GC, Houck JM, Field CA. The language of motivational interviewing and feedback: Counselor language, client language, and client drinking outcomes. Psychology of Addictive Behaviors. 2010; 24(2):190–197. [PubMed: 20565145]
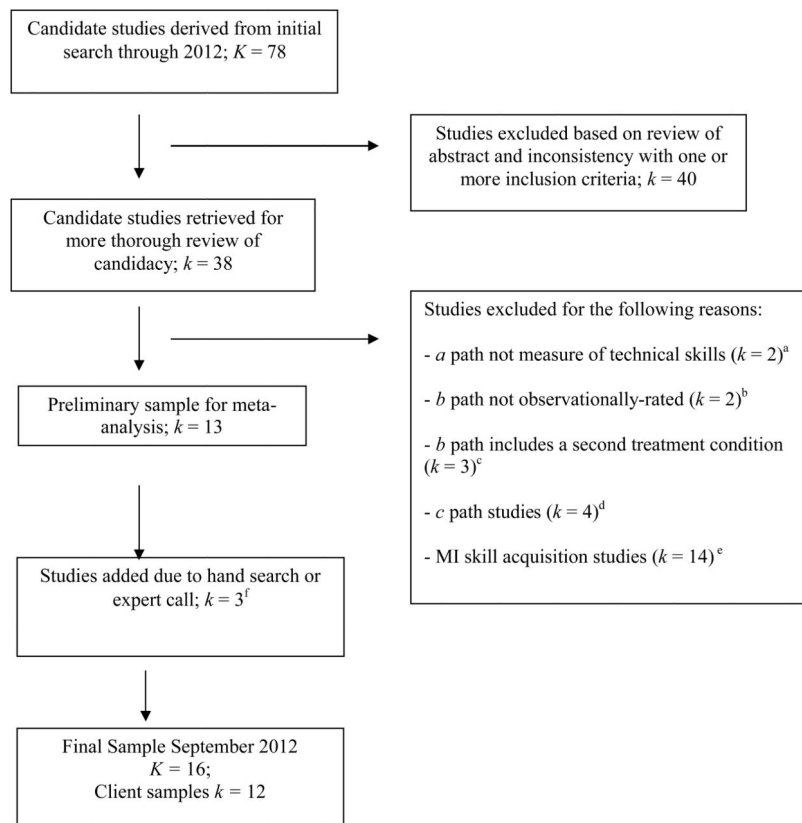
Figure 1.
Flow of primary study inclusion

Notes. K/k is defined as number of groups. [a]Glynn & Moyers, 2010; Morgenstern et al, 2013. [b]Deappen et al., 2010; Strang & McCambridge, 2004. [c]Karno et al., 2010; Moyers et al., 2007; Walker et al., 2011. [d]*c* path studies examined the effect of MI-Consistent or Inconsistent Skills on patient outcome (Gaume et al., 2009; McCambridge et al., 2011; Tollison et al., 2008/2010). [e]These studies examined the effect of MI training on technical skill acquisition. [f]Apodaca et al., 2013; Miller et al., 1993; Morgenstern et al., 2013. A total of 16 published reports tested the *a* and/or *b* paths for 12 clinical trials.

Figure 2.
Meta-analytic results on the Technical Hypothesis of MI efficacy
Notes. *** *p* < .001; ** *p* < .005; * *p* < .05 †*p* < .10.

Table 1
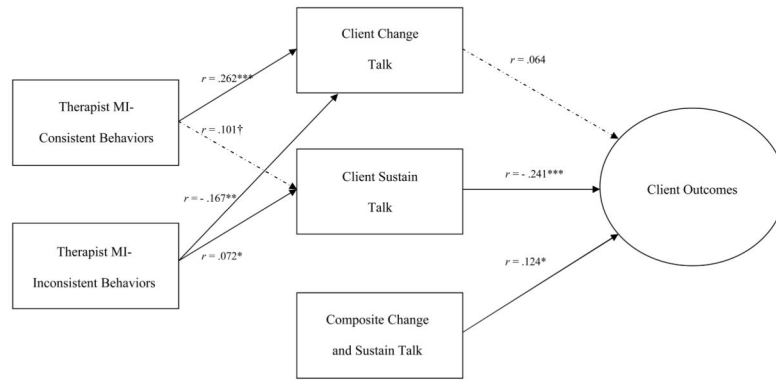
Summary and reliability data on primary study descriptors

Variable	Mean (*SD*)	Percent(*k*)	ICC[a]	Kappa[b]
Demographic factors				
Age	35.5(9.9)		.71	
Adolescent sample		8.3(1)		1.0
College/young adult sample		16.7(2)		
Adult sample		75.0(9)		
Percent female in sample	40.3(18.4)		1.0	
Majority Caucasian sample		75.0(6)		.51
Majority other ethnicity sample		25.0(2)[c]		
Diagnostic factors				
Treatment seeking sample		53.8(7)		.85
Non-treatment seeking sample		46.2(5)		
Non-dependence/abuse		58.3(7)		.32
Dependence criteria		41.7(5)		
Alcohol study		66.7(8)		1.0
Illicit drug study		16.7(2)		
Other behavior study		16.7(2)		
Intervention factors				
Session time in minutes	47.3(19.6)		.88	
Session protocol intervention		50.0(6)		.84
Manualized intervention		50.0(6)		
Supervision with session review		75.0(9)[d]		.55

Notes. K/k is defined as number of groups.

[a]Intraclass Correlation Coefficient; two-way mixed single measure. Cicchetti (1994) identifies interpretation guidelines as follows: below .40 = poor; .40 to .59 = fair; .60 to .74 = good; and .75 to 1.0 = excellent.

[b]Cohen's unweighted kappa; Landis and Koch (1977) proposed the following standards for agreement: 0 = poor, .01–.20 = slight, .21–.40 = fair, .41–.60 = moderate, .61–.80 = substantial, and .81–1 = almost perfect.

[c]Four studies had no available ethnic/racial data (Campbell et al., 2010; Hodgins et al., 2009; Miller et al., 1993; Moyers et al., 2009).

[d]Three studies reported no supervision protocol (Catley et al., 2006; Hodgins et al., 2009; Miller et al., 1993).

Table 2

a and b path studies of the MI Technical Hypothesis – Change Talk

First Author (date)	n	a Path		Target Behavior	b Path		
		Sequential Analysis	Effect Size (r)		Covariate Adjusted	Follow-up Time point	Effect Size (r)
Apodaca (2013)[a]							
MICO to CT	195	no	.18[**]	alcohol	-	-	-
MIIN to CT		no	−.21[**]				
Baer (2008)	54	-	-	drug	no	1mo	.03[b]
Catley (2006)							
MICO to CT	86	no	.84[***]	smoking	-	-	-
MIIN to CT		no	−.02[c]				
Gaume (08a/Bertholet 10)							
MICO to CT	97	yes	.10[***]	alcohol	no	12mo	.07
MIIN to CT		yes	−.16[*]				
Gaume (2010/2013)							
MICO to CT	149	yes	.15[***]	alcohol	no	6mo	−.03
MIIN to CT		yes	−.43[**]				
Hodgins (2009)	39	-	-	gambling	no	3mo	.39[*]
Miller (1993)							
MICO to CT	42	no	.37[*]	alcohol	-	-	-
MIIN to CT			−.37[*]				
Moyers (2007/2009)							
MICO to CT	63	yes	.03	alcohol	no	5wk	.20[*]
MIIN to CT		yes	−.08[*]				
Vader (2010)[d]							
MICO to CT – MIF	30	no	.71[***]	alcohol	no	3mo	.16
MICO to CT – MIO	30	no	.55[**]				−.46[*]

Notes.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

***
$p < .001$;

**
$p < .005$;

*
$p < .05$;

†
$p < .10$.

mo = month; wk = week.

[a] Apodaca and colleagues (2013) included two experimental conditions: (a) MI conducted with the participant alone and (b) MI conducted with the participant and a significant other in the session; MI with significant other was considered ineligible for the present study.

[b] Baer and colleagues (2008) only reported data for sub-code change talk indicators; these measures were pooled to the study-level.

[c] The summary score for MIIN measured at "fair" agreement (Landis & Koch. 1977).

[d] MIF = MI with Feedback. MIO = MI-Only (no Feedback); MIIN paths did not achieve acceptable inter-rater reliability and was not included in the present review.

Table 3

a and b path studies of the MI Technical Hypothesis – Sustain Talk

First Author(date)	n	a Path			b Path		
		Sequential Analysis	Effect Size (r)	Target Behavior	Covariate Adjusted	Follow-up Time point	Effect Size (r)
Apodaca (2013)[a]							
MICO to ST	195	no	−.02	alcohol	-	-	-
MIIN to ST		no	.07				
Baer (2008)	54	-	-	drug	no	1mo	−.23[†b]
Campbell (2010)	28	-	-	alcohol	no	6mo	−.42*
Catley (2006)							
MICO to ST	86	no	.46***	nicotine	-	-	-
MIIN to ST		no	.05[c]				
Gaume (08a/Bertholet 10)							
MICO to ST	97	yes	.13***	alcohol	no	12mo	−.07
MIIN to ST		yes	−.02				
Gaume (2010/2013)							
MICO to ST	149	yes	.09***	alcohol	no	6mo	−.09
MIIN to ST		yes	.22*				
Hodgins (2009)	39	-	-	gambling	no	3mo	−.28[†]
Miller (1993)							
MICO to ST	42	no	−.12	alcohol	no	12mo	−.58**[d]
MIIN to ST		no	.12				
Moyers (2009)							
MICO to ST	63	yes	−.24***	alcohol	-	-	-
MIIN to ST		yes	.07[†]				
Vader (2010)[e]							
MICO to ST – MIF	30	no	.45*	alcohol	no	3mo	−.28
MICO to ST – MIO			.59**				−.57**

Notes.

NIH-PA Author Manuscript

***
$p < .001$;

**
$p < .005$;

*
$p < .05$;

†
$p < .10$.

mo = month; wk = week.

[a] Apodaca and colleagues (2013) included a significant other in the session for one of two experimental conditions; MI with significant other was considered ineligible for the present study.

[b] Baer and colleagues (2008) only reported data for sub-code change talk indicators; these measures were pooled to the study-level.

[c] The summary score for MIIN measured at "fair" agreement (Landis & Koch. 1977).

[d] Miller and colleagues (1993) only reported data for sub-code change talk indicators; these measures were pooled to the study-level.

[e] MIF = MI with Feedback, MIO = MI-Only (no Feedback); MIIN paths did not achieve acceptable inter-rater reliability and was not included in the present review.

Table 4

b path studies of the MI Technical Hypothesis – Composite Change Talk

First Author(date)	n	Target Behavior	Covariate Adjusted	Follow-up Time point	Effect Size (*r*)
				***b* Path**	
Amrhein (2003)	84	drug	yes	12mo	.16
Campbell (2010)	40	alcohol	no	6mo	.22
Gaume (2008b)	97	alcohol	no	12mo	.06
Gaume (2013)	127	alcohol	no	6mo	.07
Hodgins (2009)	39	gambling	no	3mo	.12
Morgenstern (2012)	89	alcohol	yes	wk1	.20[†]

Notes. ***p < .001; **p < .005; *p < .05

[†]
p < .10.

mo = month; wk = week.

[a]Effect size data for Composite Change Talk were collected from studies that reported combined positive and negative frequency or strength ratings or statements of commitment only.