**MOTIVATIONAL INTERVIEWING**

Jennifer Hettema, Julie Steele, and William R. Miller

*Department of Psychology, University of New Mexico, Albuquerque, New Mexico 87131-1161; email: jhettema@unm.edu, jmsteele@unm.edu, wrmiller@unm.edu*

**Key Words**  substance abuse, health behavior, treatment outcome, meta-analysis, counseling

**Abstract**  Motivational interviewing (MI) is a client-centered, directive therapeutic style to enhance readiness for change by helping clients explore and resolve ambivalence. An evolution of Rogers's person-centered counseling approach, MI elicits the client's own motivations for change. The rapidly growing evidence base for MI is summarized in a new meta-analysis of 72 clinical trials spanning a range of target problems. The average short-term between-group effect size of MI was 0.77, decreasing to 0.30 at follow-ups to one year. Observed effect sizes of MI were larger with ethnic minority populations, and when the practice of MI was not manual-guided. The highly variable effectiveness of MI across providers, populations, target problems, and settings suggests a need to understand and specify how MI exerts its effects. Progress toward a theory of MI is described, as is research on how clinicians develop proficiency in this method.

**CONTENTS**

INTRODUCTION ................................................. 92
META-ANALYTIC METHODS ............................................ 94
Study Identification and Coding ........................................ 94
Computing Effect Sizes .............................................. 95
Comparison of Problem Areas, Comparison Group, and Motivational Interviewing Purity .......................................... 95
Analysis of Motivational Interviewing Efficacy Across Time ............. 96
Homogeneity Analyses ................................................. 96
RESULTS ............................................................... 96
Characteristics of Included Trials ........................................ 96
TREATMENT EFFECTS OF MOTIVATIONAL INTERVIEWING .......... 99
General Observations ................................................ 99
Correlates of Effect Size .............................................. 101
Effects of Motivational Interviewing by Problem Domain ................ 102
DISCUSSION ..................................................... 103
Treatment Adherence ................................................. 103
Immediacy of Effect ................................................. 104
Are Manuals a Good Idea? ........................................... 104
Matching Indications ................................................. 105
INTRODUCTION

Anyone who aspires to help others change will quickly discover that people are often less than “ready, willing, and able” to do so. The “able” part of this formula is comfortable territory for most cognitive-behavior therapists, who are quite prepared to help clients build self-efficacy and learn how to change through a rich armamentarium of effective coping strategies. Less familiar is the terrain of readiness. Often clients are expected to come already prepared with sufficient motivation for change. In substance abuse treatment, it was once common to tell less-motivated clients, “Come back when you’re ready.”

Yet, hesitancy about change is human nature. To be sure, clients present with a wide range of readiness. Some do come already convinced that something has to change. Others come reluctantly or grudgingly, nudged through the door by loved ones or the courts. It is a safe assumption that most clients seeking treatment or change are ambivalent about it: They want it, and they don’t.

Motivational interviewing (MI) was developed as a way to help people work through ambivalence and commit to change (Miller 1983). An evolution of client-centered therapy, MI combines a supportive and empathic counseling style (Rogers 1959) with a consciously directive method for resolving ambivalence in the direction of change. Drawing on Bem’s self-perception theory (Bem 1972) that people tend to become more committed to that which they hear themselves defend, MI explores the client’s own arguments for change. The interviewer seeks to evoke this “change talk”—expressions of the client’s desire, ability, reasons, and need for change—and responds with reflective listening. Clients thus hear themselves explaining their own motivations for change, and hear them reflected again by the counselor. Furthermore, the counselor offers periodic summaries of change talk that the client has offered, a kind of bouquet composed of the client’s own self-motivational statements (Miller & Rollnick 2002).

The net effect of evoking change talk in an empathic and supportive manner is to strengthen the client’s commitment to change. Verbalized intention results in an increased probability of behavior change, particularly when it is combined with a specific plan for implementation (Gollwitzer 1999). In psycholinguistic analyses of MI sessions with drug dependent people, we found that the strength of commitment language predicted drug abstinence. Stated desire, ability, reasons, and need for change all contributed to subsequent strength of commitment language, but only commitment directly predicted behavior change (Amrhein et al. 2003). To say that one wants to, can, has cause to, or needs to change is not the same as making a commitment or stating the intention to change. MI is therefore differentiated into two phases: the first is focused on increasing motivation for change, and the second on consolidating commitment (Miller & Rollnick 2002).
MI is normally brief, provided in one to two sessions. It can be delivered as a freestanding intervention, or as a motivational prelude to other treatment. It has also been common to combine the clinical method of motivational interviewing with other intervention components, which have been called adaptations of MI (AMIs) (Burke et al. 2003). The most widely used AMI is motivational enhancement therapy (MET), which combines MI with personal feedback of assessment results (Miller et al. 1992).

Like other psychotherapies, MI is a complex and skillful method that is learned over time. Counselors sometimes come to MI workshops expecting to learn tricks for getting people to do what counselors want them to do. On the contrary, MI is a systematic and collaborative method for helping people to explore their own values and motivations, and how these may be served by status quo or behavior change. It emphasizes and honors client autonomy, to choose whether, when and how to change. When done well, MI involves listening more than telling. It does not operate from a deficiency model that seeks to instill knowledge, insight, skills, correct thinking, or even motivation. Rather, the counselor seeks to evoke the client’s own motivation, with confidence in the human desire and capacity to grow in positive directions. Instead of implying that “I have what you need,” MI communicates, “You have what you need.” In this way, MI falls squarely within the humanistic “third force” in the history of psychotherapy. Nevertheless, MI is compatible with a variety of other approaches and appears to amplify the efficacy of treatment methods with which it is combined.

Proficiency in MI is not readily acquired by reading about it, viewing videotapes, or attending a clinical workshop (Miller & Mount 2001). Proper training focuses instead on helping clinicians learn how to learn MI from their clients. Once counselors learn to recognize and evoke change talk and committing language, clients thereafter provide continuous and immediate in-session reinforcement for good practice. Client resistance, on the other hand, represents immediate feedback of dissonance and serves as a cue to shift strategies. Within MI, “resistance” is simply client speech that defends and expresses commitment to status quo; in other words, it reflects the other side of the client’s ambivalence. Pushing against resistance tends to focus on and amplify it. Instead, the interviewer acknowledges and rolls with resistance, calling attention to both sides of the ambivalence and redirecting the emphasis toward change.

MI differs from client-centered counseling in its directive intention. Some have maintained that Rogers himself was unconsciously directive, differentially attending to and reinforcing certain types of client speech (Truax 1966). In MI, such differential response to change talk is conscious and strategic. This means, of course, that MI is appropriate when there is a clear desired direction for change. That direction may come from the client’s own expressed desires, from the counselor’s perspective, or from the context within which counseling occurs. Interesting ethical dilemmas can arise when therapists and clients disagree on the perception of a problem and the need to change. MI has been argued to lie on a continuum between passivity and coercion and seeks to resolve mismatches between clients.
and counselors by evoking the clients’ intrinsic motivations (Miller 1994, Miller & Rollnick 2002).

Research indicates that MI is particularly useful with clients who are less motivated or ready for change, and who are more angry or oppositional. For these populations, action-oriented counseling with a goal of behavior change is likely to evoke resistance and reactance. From a transtheoretical perspective, this happens because of a mismatch in stages of change: The counselor is working at the action stage, whereas the client is in the earlier precontemplation or contemplation stage (Prochaska & DiClemente 1984). In the case of clients who are less ready for change, MI meets them where they are and invites them to move along through contemplation, preparation, and action. For clients who indicate readiness to change, MI may be less useful, and some findings indicate that it can be counterproductive. If such clients subsequently show ambivalence in action-oriented counseling, one can always fall back to an MI style.

The treatment outcome literature for MI is growing rapidly and has spread well beyond its original focus on addictive behaviors. Our primary purpose in this chapter is to provide an up-to-date summary of the evidence base for MI, drawing data primarily, but not exclusively, from controlled clinical trials. The findings that we summarize here are based on a new meta-analysis, the full scope of which is beyond the space limitations of this chapter. Full details of the meta-analysis and a comprehensive bibliography of MI are available at http://www.motivationalinterview.org/.

META-ANALYTIC METHODS

Study Identification and Coding

In order to identify MI treatment outcome studies, we searched PsycINFO using the term “motivational interviewing,” and hand-searched bibliographies from the motivational interviewing web page (http://www.motivationalinterview.org/) and previous reviews (Burke et al. 2003, Dunn et al. 2001, Miller & Wilbourne 2003). Studies having (a) at least one group or individual intervention with components of MI, and (b) at least one posttreatment outcome measure were included in the overall pool for analyses investigating within-group effect sizes. In addition, studies contributing between-group effect sizes required (c) at least one control condition or comparison intervention without any components of MI, and (d) a procedure to provide pretreatment equivalence of groups (e.g., randomization, cohort, or sequential group assignment).

All outcome studies were independently coded by the first two authors (J. Hettema and J. Steele). The characteristics of included studies (type, goal, format, setting, intervention agent, treatment components, and sample characteristics) were categorized using a coding manual from prior treatment outcome reviews (Miller & Wilbourne 2003), with adaptations for the specific content of MI. Classification discrepancies were resolved by consensus of the coders, with reference to the original article and coding manual. All studies were also rated using 12 methodological
quality criteria from the same coding system, including method for assignment to groups, presence of quality control of treatment, follow-up rate, follow-up duration, type of follow-up data collection, collateral verification of self-report, objective verification of follow-up data, inclusion of treatment dropouts in analyses, consideration of cases lost to follow-up, masked follow-up data collection, acceptable statistical analyses, and the inclusion of multiple sites. Total methodological quality scores were computed, with a possible range from 0 to 16. In addition, we coded information on the amount and type of MI training provided to interventionists, and specific components of MI reported to have been included in the interventions.

Computing Effect Sizes

For each study, effect sizes and confidence intervals were computed for all outcome variables related to the target problem, and for which sufficient information was provided. As feasible, study authors were contacted for missing information. When no other option was available, effect sizes reported in previously published meta-analyses were used (Bien et al. 1993, Burke et al. 2003, Dunn et al. 2001). When insufficient information was provided to determine effect sizes and significance tests indicated \( p > 0.05 \), zero effect sizes were assigned.

When calculating within-group effect sizes, baseline mean values of all included variables were compared to every follow-up point. For between-group calculations, mean MI scores on every included variable were compared to every other investigated treatment condition at all follow-up points. When mean, standard deviation, and sample size information were reported, an unbiased estimator of effect size \( (g) \) was calculated using the following formula (Hedges & Olkin 1985):

\[
g = \frac{J(N-2)}{2} \times \left( \frac{\bar{Y}_E - \bar{Y}_C}{s} \right)
\]

where \( J(N-2) \) is a bias correction factor, \( \bar{Y}_E \) and \( \bar{Y}_C \) are the experimental and control group means, and \( s \) is the pooled sample standard deviation. When mean, standard deviation, or sample size information was not provided, effect sizes were estimated from significance tests. \( F, t, \) or chi-square statistics were transformed to \( r \) values and then converted to effect sizes \( (d) \) using the following formula (Rosenthal 1991):

\[
d = \frac{2r}{\sqrt{1 - r^2}}
\]

For all effect sizes, 95% confidence intervals were then calculated using the following formula (Hedges & Olkin 1985, p. 86):

\[
\sigma^2(d) = \left( \frac{n^E + n^C}{n^E n^C} \right)\left( \frac{d^2}{2(n^E + n^C)} \right)
\]

In addition, we calculated for each study a combined effect size \( (d_c) \), averaging all variables at each follow-up point using weighted linear combinations (Hedges & Olkin 1985, pp. 109–117). To minimize the variance of the combined effect sizes, weights that were inversely proportional to the variance of each effect size were assigned to each variable included in the analyses.

Comparison of Problem Areas, Comparison Group, and Motivational Interviewing Purity

This review includes studies across all behavior domains for which the efficacy of MI has been investigated, and we report effect sizes by target behaviors. We further differentiated trials comparing MI to untreated control groups from those in which
MI was added to or compared with other types of active treatment. A previous meta-analysis reported slightly larger effects of MI when added to other treatment than when tested as a stand-alone intervention (Burke et al. 2003). Finally, we did our best to differentiate studies of “pure” MI from those in which MI was combined with another established treatment. We computed composite effect sizes to address each of these issues, using the combined effect size from each relevant study to determine the relative efficacy of MI across problem areas, design types, and in studies with more “pure” forms of MI versus those in which MI was combined with another treatment. In all, we estimated more than 884 effect sizes in preparing this review.

Analysis of Motivational Interviewing Efficacy Across Time

Most studies of MI have reported outcome data across several follow-up points. To provide cross-study consistency, we classified follow-ups as having occurred at posttreatment and at the following posttreatment intervals: 1–3 months, 4–6 months, 7–12 months, 13–24 months, and longer than 2 years. Combined between-group effect sizes were calculated for all data during each of the follow-up intervals. In addition, combined within-group effect sizes for MI were calculated for each time interval, comparing each follow-up variable value with its baseline level.

Homogeneity Analyses

To determine the appropriateness of later statistical procedures, such as t-tests and multiple regression analyses, homogeneity analyses were conducted on groups of effect sizes that were entered into these analyses. T-tests and multiple regression analyses assume homoscedasticity, or that nonsystematic variance is equal across observations, and little is known about the violation of this assumption on these conventional statistical methods (Hedges & Olkin 1985). A Q statistic was calculated and tested for significance for each group that would be entered into a later analysis. A significant Q statistic indicates that the group is statistically heterogeneous.

RESULTS

Characteristics of Included Trials

STUDY DESIGN For full details of the characteristics of each trial, see the Supplemental Material link for Supplemental Table 1 in the online version of this chapter or at http://www.annualreviews.org/. Seventy-two studies met inclusion criteria for this meta-analysis. The studies tested the efficacy of motivational interviewing within the following behavioral domains: alcohol (31), smoking (6), HIV/AIDS (5), drug abuse (14), treatment compliance (5), gambling (1), intimate relationships (1), water purification/safety (4), eating disorders (1), and diet and exercise (4).
<table>
<thead>
<tr>
<th>Target Problem</th>
<th>Combined $d_c$ across all follow-up points</th>
<th>Combined $d_c$ at follow-ups ≤ three months</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All studies Untreated Additive Treatment</td>
<td>All studies Untreated Additive Treatment</td>
</tr>
<tr>
<td>Alcohol</td>
<td>0.26 (N = 31) 0.38 (N = 14) 0.33 (N = 5) 0.11 (N = 13)</td>
<td>0.41 (N = 18) 0.44 (N = 9) 0.28 (N = 3) 0.38 (N = 6)</td>
</tr>
<tr>
<td></td>
<td>(0.18, 0.33) (0.20, 0.56) (0.23, 0.44) (0.05, 0.17)</td>
<td>(0.31, 0.51) (0.30, 0.59) (0.03, 0.54) (0.23, 0.53)</td>
</tr>
<tr>
<td>Smoking</td>
<td>0.14 (N = 6) 0.13 (N = 2) 0.17 (N = 5)</td>
<td>0.04 (N = 2) 0.01 (N = 1) 0.05 (N = 1)</td>
</tr>
<tr>
<td></td>
<td>(0.09, 0.20) (0.04, 0.22) (0.08, 0.25)</td>
<td>(−0.08, 0.16) (−0.27, 0.30) (−0.09, 0.18)</td>
</tr>
<tr>
<td>HIV</td>
<td>0.53 (N = 5) 0.12 (N = 3) 0.94 (N = 2)</td>
<td>0.71 (N = 4) 0.12 (N = 3) 3.4 (N = 1)</td>
</tr>
<tr>
<td></td>
<td>(0.24, 0.81) (−0.04, 0.28) (0.41, 1.46)</td>
<td>(0.24, 1.19) (−0.04, 0.28) (2.4, 4.5)</td>
</tr>
<tr>
<td>Drugs</td>
<td>0.29 (N = 13) 0.45 (N = 6) 0.53 (N = 2) 0.12 (N = 6)</td>
<td>0.51 (N = 9) 0.69 (N = 5) 0.53 (N = 2) 0.02 (N = 2)</td>
</tr>
<tr>
<td></td>
<td>(0.15, 0.43) (0.16, 0.74) (−0.05, 1.12) (−0.05, 0.20)</td>
<td>(0.13, 0.90) (0.05, 1.32) (−0.05, 1.12) (−0.08, 0.13)</td>
</tr>
<tr>
<td>Treatment compliance</td>
<td>0.72 (N = 5) 0.10 (N = 2) 0.80 (N = 3)</td>
<td>0.42 (N = 4) 0.16 (N = 2) 0.75 (N = 2)</td>
</tr>
<tr>
<td></td>
<td>(0.56, 0.89) (−0.15, 0.36) (0.64, 0.97)</td>
<td>(0.21, 0.63) (0.16, 0.48) (0.41, 1.09)</td>
</tr>
<tr>
<td>Gambling</td>
<td>0.29 (N = 1) 0.46 (N = 1) 0.24 (N = 1)</td>
<td>0.44 (N = 1) 0.46 (N = 1) 0.43 (N = 1)</td>
</tr>
<tr>
<td></td>
<td>(0.16, 0.42) (0.17, 0.74) (0.09, 0.40)</td>
<td>(0.27, 0.61) (0.17, 0.74) (0.21, 0.65)</td>
</tr>
<tr>
<td>Water purification/safety</td>
<td>0.30 (N = 4) 0.30 (N = 4) 0.51 (N = 2)</td>
<td>0.51 (N = 2)</td>
</tr>
<tr>
<td></td>
<td>(0.05, 0.55) (0.05, 0.55) (0.30, 0.72)</td>
<td>(0.30, 0.72)</td>
</tr>
<tr>
<td>Eating disorder</td>
<td>−0.07 (N = 1)</td>
<td>−0.07 (N = 1)</td>
</tr>
<tr>
<td></td>
<td>(−0.42, 0.26)</td>
<td>(−0.42, 0.26)</td>
</tr>
<tr>
<td>Diet and exercise</td>
<td>0.78 (N = 4) 0.78 (N = 4) 0.14 (N = 1)</td>
<td>0.14 (N = 1)</td>
</tr>
<tr>
<td></td>
<td>(0.41, 1.16) (0.41, 1.16) (−0.16, 0.44)</td>
<td>(−0.16, 0.44)</td>
</tr>
</tbody>
</table>

*Note: $p < 0.05$; N represents the number of studies from which variables were taken (some studies had outcome variables in two comparison categories).
In the analyzed studies, MI was seldom given alone, but was typically combined with feedback and often some other form of treatment. In 41 studies, treatment groups received MI or MI plus feedback only, whereas in 31 studies, MI was combined with some other type of intervention, including education, self-help manuals, relapse prevention, cognitive therapy, skills training, Alcoholics Anonymous, stress management, and treatment as usual for the particular setting. Comparison groups also differed widely across studies. In 21 studies, MI was compared to a no-treatment or placebo condition. Five studies investigated the additive effects of MI to standard treatment, whereas six studies directly contrasted MI with an unspecified standard treatment. Seven studies investigated the effects of MI when added to another established treatment, twenty-five studies contrasted MI with another established treatment, six studies had mixed designs, and two studies solely investigated within-group change.

As discussed above, all studies were coded for 12 dimensions of methodological quality, yielding methodological quality scores that ranged from 4 to 16 (mean = 10.76, SD = 2.43), slightly higher than the mean score (10.68) reported for 361 alcoholism clinical trials in general (Miller & Wilbourne 2003). In comparison to these 361 trials, studies of MI were more likely to report some form of intervention quality control (78% versus 57%) and to be multisite trials (28% versus 5%), but were less likely to follow clients for 12 months or longer (18% versus 51%) or to complete follow-up with 70% or more of enrolled participants (45% versus 75%). The duration of follow-up ranged from 0 to 60 months posttreatment (mean = 8.8, SD = 10.28).

All outcome variables for which effect sizes could be calculated were enumerated for each study. The number of reported outcome variables ranged from 1 to 12 (mean = 3.3, SD = 2.3). To avoid capitalization on chance by the number of statistical tests conducted, we combined effect sizes across all reported outcome variables in each study.

CHARACTERISTICS OF MOTIVATIONAL INTERVIEWING For full details of the characteristics of MI for each trial see the Supplemental Material link for Supplemental Table 2 in the online version of this chapter or at http://www.annualreviews.org/. Characteristics of MI were also coded for all studies. As a rough index of the degree to which each study had implemented MI, we coded whether interventions were specified as including the following components of MI: being collaborative, being client centered, being nonjudgmental, building trust, reducing resistance, increasing readiness to change, increasing self-efficacy, increasing perceived discrepancy, engaging in reflective listening, eliciting change talk, exploring ambivalence, and listening empathically. The total number of these strategies reported to have been implemented in interventions identified as MI ranged from 0 to 12 (mean = 3.6, SD = 2.8).

The duration of MI interventions also varied. In 68 studies that reported these data, MI duration ranged from 15 minutes to 12 hours, with an average dose of about two sessions (mean = 2.24 hours, SD = 2.15). When MI was combined with
other treatment components, “duration” included only the time committed to MI. Comparison group treatment durations ranged from 0 to 28 hours (mean = 2.89, SD = 5.57). The difference in treatment duration between MI and comparison groups ranged from −25 hours (the comparison treatment was 25 hours longer than MI) to +6 hours [MI was 6 hours longer than the no-treatment control (mean = −0.48, SD = 4.9)].

Of the 72 studies included in the analyses, most (74%) reported that the MI intervention had been standardized by a manual or a specific training. For 13 studies that reported amount of training time, a mean of 9.92 (SD = 7.35) hours was spent in training. Only 26 studies (29%) provided any kind of posttraining support (such as supervision) for therapists, and only 21 studies (36%) included any form of competency or fidelity assessment after initial training.

MI was delivered in a variety of settings, including aftercare/outpatient clinics (15), inpatient facilities (11), educational settings (6), community organizations (6), general practitioner offices (5), prenatal clinics (3), emergency rooms (2), employee assistance programs (2), halfway houses (2), over the telephone (3), in patients’ homes (1), in jail (1), in mixed settings (7), or in unspecified treatment settings (8). The agents implementing the MI, when specified, included paraprofessionals or students (8), master’s level counselors (6), psychologists (6), nurses (3), physicians (2), dieticians (1), and medically a mix of varying levels of professionals (22).

CHARACTERISTICS OF STUDY SAMPLES For full details of the characteristics of each trial sample see the Supplemental Material link for Supplemental Table 3 in the online version of this chapter or at http://www.annualreviews.org/. The 72 studies enrolled between 21 and 952 participants (mean = 198.16, SD = 204.39), for a total of 14,267 clients. On average, the samples included 54.77% males (range: 0%–100%), and ranged in age from 16 to 62 (mean = 34.11, SD = 8.96). Only 37 studies specified ethnic composition, of which 16 samples (43%) were comprised primarily of participants from U.S. minority groups, including 10 with predominantly or entirely African American samples. Problem severity varied widely, and eight samples specifically recruited participants with concomitant substance use and mental disorders.

TREATMENT EFFECTS OF MOTIVATIONAL INTERVIEWSING

General Observations

Before examining MI effects by target problem areas, we offer some broad observations from our analysis of 72 outcome studies. The effect sizes for each variable at every follow-up point is available at the Supplemental Material link for Supplemental Table 4 in the online version of this chapter or at http://www.annualreviews.org/
First of these observations is the wide variability in effect sizes across studies, even within problem areas. In studies of alcohol abuse, for example, although most trials have reported statistically significant effects of MI, the observed effect sizes have varied from \( d_c = 0 \) to more than 3.0 (where \( d_c = 1.0 \) represents a between-group difference of one standard deviation). This means that in using ostensibly the same treatment method (MI) with the same target problem, very different effects are obtained across sites and populations. In Project MATCH, a nine-site study of treatments for alcohol use disorders, the relative efficacy of an MI-based intervention varied significantly across sites and therapists despite extensive efforts to standardize training and treatment procedures (Project MATCH Research Group 1998). Thus, it appears that variation in the delivery of MI can have substantial impact on its outcome.

A second broad observation is that an effect of MI tends to be seen early and to diminish across a year of follow-up. To examine this, we combined effects for all variables from all studies within specific follow-up period ranges. As displayed in Figure 1, relative effect sizes for MI decrease across time. Across all studies, \( d_c \) was 0.77 (95% confidence interval: 0.35, 1.19) at 0 to 1 month posttreatment, 0.39 (0.27, 0.50) at >1 to 3 months, 0.31 (0.23, 0.38) at >3 to 6 months, 0.30 (0.16, 0.43) at >6 to 12 months, and 0.11 (0.06, 0.17) at follow-ups longer than 12 months. An interesting exception to this trend, seen in Figure 1, is found in studies where the additive effect of MI is tested. In these studies, clients are typically randomized to receive or not to receive MI at the beginning of a standard or specified treatment

![Figure 1](image-url)  
**Figure 1** Combined effect sizes of motivational interviewing across follow-up intervals.
program. In this case, the effect of MI in improving outcome is maintained or increased over time, hovering around $d_c = 0.60$.

Outcome variability, however, makes it difficult to specify a meaningful average effect size for MI without regard to problem domain, population, interventionists, or follow-up duration. A full table of combined between-group effect sizes for each included study can be viewed online. See the Supplemental Material link in the online version of this chapter or at http://www.annualreviews.org/. The combined effect sizes (pooling across outcome variables and follow-up points) for individual studies ranged from $-0.19$ to $3.25$ (mean $= 0.43$, SD $= 0.62$). Using $d_c$ for all reported outcome variables across all follow-up points, 38 of the studies (53%) showed a significant effect favoring MI ($p < 0.05$).

Correlates of Effect Size

**STUDY CHARACTERISTICS** We also examined relationships between observed combined effect size ($d_c$) and a number of study attributes as potential moderators of outcome. In regression as well as correlational analyses, we found no significant relationship between $d_c$ and study characteristics including methodological quality, number of outcome variables, longest follow-up point, MI purity, type of comparison group, or problem area.

**MI CHARACTERISTICS** In multiple regression analyses, we found that $d_c$ was not significantly predicted by our measures of MI duration, purity, counselor training, or posttraining support. Of MI delivery characteristics, only the presence of a manual was significantly related to outcome, predicting $8.5\%$ of the variance in $d_c$ ($\beta = -0.292$, $p < 0.05$). The direction of this difference was such that studies not reporting use of a manual had a mean $d_c = 0.65$ (SD $= 0.62$), whereas those standardizing treatment with a manual reported a mean $d_c = 0.37$ (SD $= 0.62$). A follow-up independent sample t-test reflected this difference as a trend ($t = 1.53$, $p = 0.28$). It should be noted that no studies provided data allowing for within-study comparison of manual-guided versus nonmanual-guided MI. Because the evidence that manual-guided treatments are associated with smaller effect size comes solely from between-study comparisons, it is possible that other important differences between studies exist.

**SAMPLE CHARACTERISTICS** Similarly, we regressed $d_c$ onto study sample characteristics including mean age, gender composition, ethnic composition, and problem severity. Only ethnic composition significantly predicted $d_c$, accounting for $19\%$ of variance ($\beta = 0.434$, $p < 0.05$). A follow-up test ($t = -0.39$, $p < 0.05$) revealed that effects of MI were significantly larger for minority samples ($M d_c = 0.79$) than for non-minority white samples ($M d_c = 0.26$).

**OUTCOME MEASURES** Within behavioral domains, studies utilized a wide variety of treatment outcome measures. Although most behavioral domains had too
few studies and too many different outcome variables to form meaningful groups, alcohol outcome variables could be divided into quantity, frequency, intoxication (blood alcohol concentration, or BAC) level, and alcohol-related problems categories. Combined effect sizes were determined for each of these variables across studies and follow-up points. A $d_c = 0.30 (0.09, 0.52; p < 0.05)$ was found for quantity variables, $d_c = 0.31 (0.18, 0.44; p < 0.05)$ for frequency variables, $d_c = 0.22 (0.10, 0.34; p < 0.05)$ for BAC variables, and $d_c = 0.08 (−0.02, 0.19; p > 0.05)$ for alcohol-related problems. For smoking studies, a $d_c = 0.15 (−0.06, 0.23; p < 0.05)$ was found for abstinence outcome variables, and $d_c = 0.11 (0.00, 0.21; p > 0.05)$ for quit attempt variables. Variables from HIV studies could be divided into knowledge with $d_c = 1.46 (−0.54, 3.45; p > 0.05)$, behavioral intentions with $d_c = 0.88 (0.05, 1.72; p < 0.05)$, and sexual risk behaviors with $d_c = 0.07 (−0.05, 0.19; p > 0.05)$.

Effects of Motivational Interviewing by Problem Domain

Table 1 provides a concise summary of effect sizes, combined across outcome variables, for studies of MI in various problem domains. In contrast to the above-reported analyses (Figure 1), which showed substantial reduction in $d_c$ over time, Table 1 provides separate $d_c$ means in the short-term (up to three-month follow-up), and then combined across all follow-up points. Combined effect sizes are further subdivided based on the nature of the comparison group: (a) MI versus no treatment or placebo, (b) MI versus no MI added to standard or specified treatment, or (c) MI contrasted with a standard or specified treatment. For studies with mixed comparisons, individual variables were selected based on comparison type, and were categorized appropriately.

ADDICTIVE BEHAVIORS In terms of volume of studies, the strongest support by far for MI efficacy is in the area for which it was originally designed: altering substance use (Miller 1983). A total of 32 trials have focused on alcohol abuse, yielding $d_c$ values ranging from $−0.08$ to $3.07$, with a mean of $0.41$ posttreatment, and $0.26$ across all follow-up points. The largest effects (all $>0.7$) were reported in studies comparing MI with no treatment (Gentilello et al. 1999), a wait-list control (Kelly et al. 2000) or education (Graeber et al. 2003), or adding MI to standard treatment (Aubrey 1998, Brown & Miller 1993). An additional 13 trials tested the between-group effect of MI in addressing illicit drug use, again with a large range of effects (0 to 1.81). Here effect sizes on average were larger at early than at later follow-ups (0.51 versus 0.29). Curiously, MI appears to have been largely unsuccessful to date in promoting smoking cessation. Six MI trials yielded only one small effect collapsing across outcome variables (Butler et al. 1999). We are aware, however, of several unpublished positive trials that may soon alter this picture with regard to smoking. One study reported significant effects of MI in treating pathological gambling (Hodgins et al. 2001).
HEALTH BEHAVIORS  MI has also been tested with other health behaviors in the context of health promotion (Miller 2004). Large but inconsistent effects ($d$, from $-0.19$ to $3.25$) have been reported in five trials of MI for HIV risk reduction. Thevos and colleagues have reported large effects of MI to encourage the adoption of water purification/safety technology in rural African villages (Thevos et al. 2000, 2002/2003). Encouraging effects have also been reported for MI in promoting adherence to diet and exercise programs. A single study found no difference between MI and brief behavior therapy in treating bulimia.

TREATMENT ADHERENCE  Finally, several studies have reported large effects of MI in promoting treatment engagement, retention, and adherence. As noted above, the effects of MI appear to persist or increase over time when added to an active treatment.

DISCUSSION

Across a growing array of problem areas, MI generally shows small to medium effects in improving health outcomes. As a stand-alone brief intervention, MI has been particularly well tested and found promising in addressing addictive behaviors, with the notable exception (to date) of smoking cessation. Further research is needed to determine the reliability of and possible explanations for the discrepant findings observed for smoking behaviors. Applications to health behavior, particularly in the management of chronic illnesses, have been expanding rapidly, and initial trials suggest similar benefit to that observed with addictive behaviors.

It is clear, however, that MI as practiced in trials to date does not consistently improve outcome. Even among studies focused on the same problem domain, high variability exists in effects across studies and therapists.

An obvious research direction, therefore, is to identify factors that influence the effectiveness of MI, including specific factors that mediate and moderate its effects. With a reasonable base of clinical trials supporting specific efficacy, research has recently turned to a search for “active ingredients” and aspects of MI delivery that influence outcomes. This search has been impeded, however, because few studies have detailed how interventionists were trained, provided documentation of the fidelity of delivery of MI, or included process measures to relate to outcomes (Burke et al. 2002). In some cases (e.g., Kuchipudi et al. 1990), the brief descriptions of treatment delivered as MI appear to be inconsistent with the spirit and principles described by its progenitors (Rollnick & Miller 1995). Progress toward a theory of MI efficacy is briefly discussed in the final section of this chapter.

Treatment Adherence

Several trends emerged from our meta-analysis. One is that relatively high effect sizes are often observed when MI is added at the outset of a treatment program,
including unspecified “treatment as usual” (Aubrey 1998, Brown & Miller 1993, Daley et al. 1998). This is somewhat counterintuitive, in that larger effect sizes might be expected when MI is compared with no treatment, rather than having to exert an additive effect above active treatment. Significant improvement in treatment outcome when MI is added appears to be attributable to its effects on treatment retention and adherence. In a randomized trial, Brown & Miller (1993) found that therapists in an inpatient substance abuse treatment program who were unaware of which patients had received MI, reliably rated the MI group as more motivated and adherent and as having better prognosis. These therapist ratings, in turn, mediated the effect of MI in doubling posttreatment abstinence rates. Large effects are also reported when treatment retention and adherence are the specific targets of MI. Aubrey (1998) reported a doubling of outpatient substance abuse treatment sessions attended by adolescents given a single session of MI at intake, as well as a doubling of three-month abstinence rates.

Immediacy of Effect

Controlled trials also commonly report a rapid impact of MI, with a gradual decrease of effect size across time. This is, of course, a common finding for discrete interventions. During eight weeks of drug administration, for example, a medication may yield significant benefits that subsequently fade after dosing is discontinued. In part, this decrease in between-group effects is attributable to a “catching up” of the control/comparison groups with which MI is compared. If MI is offered as a stand-alone intervention, long-term effects may be enhanced by booster sessions or stepped care. When MI is used as a prelude to treatment, however, its effects appear to endure across time, suggesting a synergistic effect of MI with other treatment procedures.

Are Manuals a Good Idea?

An unexpected finding of our meta-analysis was the relationship between effect size and the use of manuals to guide MI delivery. Our finding that manual-guided MI was associated with smaller effect sizes bears replication and further exploration.

We have had, however, one salient experience related to manual-guided MI. Following a series of findings that an early MI session improves treatment outcomes, we conducted a large randomized trial in two public substance abuse treatment programs (Miller et al. 2003). Clients were randomly assigned to receive or not to receive a single session of MI shortly after treatment intake. The MI was manual guided and participants were followed for one year. Contrary to prior trials, we found no significant benefit of MI.

Subsequent psycholinguistic analyses of these MI sessions revealed an informative pattern (Amrhein et al. 2003). Clients who subsequently abstained from drug use during follow-up had shown a characteristic pattern of increasing motivation for and commitment to abstinence over the course of the MI session. Nonresponders,
in contrast, showed a similar increase in motivation and commitment, which sud-
denly reversed in the final minutes of the session and crashed back to zero. What
happened? The treatment manual, designed to complete MI in one session, in-
structed therapists to end the session by constructing a concrete behavior change
plan regardless of whether the client seemed ready to do so. This would have the
predictable (but unanticipated) effect, during the closing minutes of the session,
of eliciting resistance from clients who were less ready for change, which in turn
would be expected to undermine behavior change. The problem, it seems, is that
the therapists did exactly what the manual instructed them to do, pressing forward
to complete the change plan even if the client resisted, which is itself a violation
of good MI practice.

Matching Indications

Another unexpected result of our meta-analysis was the finding of larger effects of
MI with U.S. samples comprised primarily or exclusively of people from ethnic
minority groups. We have no theoretical explanation for this finding, but it does
converge with a recently completed reanalysis of data from a multisite alcoholism
treatment trial (Villanueva et al. 2003). Analyzing treatment data for only Native
American participants in Project MATCH, we found significantly better outcome
for those assigned to 4-session MI (motivational enhancement therapy) than for
those assigned to 12-session cognitive-behavior therapy or 12-step facilitation
therapy. Our informal experience in MI training with Native American populations
suggests that the client-centered, supportive, and nonconfrontational style of MI
may resemble the normative communication style of Indian populations, at least in
the American Southwest, thereby representing a culturally congruent intervention.
Similar analyses, however, failed to find an advantage for MI in African American
(Tonigan et al. 2003) or Hispanic American (Arroyo et al. 2003) clients.

MI also appears to be differentially effective with clients who are more angry and
resistant, or less ready for change (Heather et al. 1996, Project MATCH Research
Group 1997). This is consistent with the original intent and theoretical rationale
for MI. Conversely, MI may be contraindicated for clients who are already clearly
committed to change and ready for action.

TOWARD A THEORY OF MOTIVATIONAL
INTERVIEWING

The high variability of effect sizes combined with the frequency of observed sig-
nificant effects indicates that MI is an active treatment, but that the mechanisms
of action are not well understood. Our crude measure of MI purity (the number
of MI-particular components mentioned in an article) failed to predict effect size.
Although there are clear therapist differences in effectiveness in delivering MI, we
have been unsuccessful in predicting MI proficiency from personal characteristics
HETTEMA ■ STEELE ■ MILLER

of counselors (Project MATCH Research Group 1998). This suggests that it may be fruitful to examine therapeutic processes occurring within MI sessions, as possible correlates of treatment outcome.

In its origins (Miller 1983), MI was not derived from theory, but rather it arose from specification of principles underlying intuitive clinical practice. The client-centered phenomenological perspective of Carl Rogers (1959), which was clearly influential as a guiding spirit of MI, emphasized empathic understanding and radical acceptance as triggers for change. Early conceptual ties were also made to cognitive dissonance (Festinger 1957) and self-perception theory (Bem 1972), based on the reasoning that when people verbally justify behavior change they are more likely to follow through with it (Miller & Rollnick 1991).

MI places strong emphasis on eliciting the client’s own perceptions, values, and motivations for change. In Socratic fashion, it should be the client rather than the counselor who makes the arguments for change. The reasoning behind this is that people in need of change, including those who present for formal treatment, are normally also ambivalent about change. A counselor who advocates for change is likely to elicit from the client the opposite (resistance) side of the client’s own ambivalence. That might be harmless enough, except for the robust finding that people tend to become more committed to positions that they defend verbally (Bem 1967). Thus, people can literally talk themselves out of (or into) behavior change.

Therefore, counselors should act in a manner that calls forth the prochange side of client ambivalence, the side that elicits the client’s own motivations for change. Conversely, counselors should assiduously avoid the position in which they argue for change while the client argues against it. MI is, in essence, both a counseling style and a set of clinical strategies and skills for evoking change talk from clients, and for defusing resistance when it arises (Miller & Rollnick 2002).

Over the two decades since MI was introduced, data have shaped an emergent theory of the inner workings of this approach. In simplest form, the theory is expressed in three hypotheses:

1. Counselors who practice MI will elicit increased levels of change talk and decreased levels of resistance from clients, relative to more overtly directive or confrontational counseling styles.
2. The extent to which clients verbalize arguments against change (resistance) during MI will be inversely related to the degree of subsequent behavior change.
3. The extent to which clients verbalize change talk (arguments for change) during MI will be directly related to the degree of subsequent behavior change.

We have found strong support for the first two of these hypotheses. MI does roughly double the rate of change talk and halve the rate of resistance, relative to action-focused counseling or confrontation (Miller et al. 1993). The counseling skill of accurate empathy (Truax & Carkhuff 1967) has been particularly linked
to improved outcomes in treating alcohol problems (Miller & Baca 1983, Miller et al. 1980, Valle 1981). We also have found that frequency of client resistance predicted continued drinking after treatment (Miller et al. 1993). Thus, client responses appear to be highly influenced by counselor style, and in turn predict treatment outcome.

We consistently failed to find support, however, for the third hypothesis—that increased client change talk would predict behavior change. Frequency of change talk statements, which we usually measured during the first 20 minutes of an MI session, was unrelated to subsequent behavioral outcomes. This obviously posed a serious problem for the fledgling theory of MI.

Collaboration with psycholinguist Paul Amrhein led to a different approach to analyzing client speech. Amrhein suggested that we had been combining too many speech events in our single concept of change talk, and recommended disaggregating it into natural language components: desire, ability, reasons, need, and commitment. He analyzed more than 100 entire MI sessions, meticulously coding each client utterance for these speech events. In addition to counting them (frequency), he also rated the strength of motivation reflected in the client’s speech. To say, “I’ll think about it,” or “I’ll try,” for example, reflects a much lower level of commitment than “I promise” or “I will.”

The results were striking (Amrhein et al. 2003). Only one of the subtypes of change talk—commitment—predicted behavior change. Furthermore, it was not the frequency but rather the strength of commitment language, and more particularly the pattern of commitment across the session, that robustly predicted behavioral outcomes, in this case, drug abstinence. Desire, ability, reasons, and need did not predict change, but all four did predict the emergence of commitment, which in turn was prognostic of change. His psycholinguistic findings gave substance to the early intuitive distinction between two phases of MI (Miller & Rollnick 1991). In phase 1 of MI, the goal is to enhance motivation for change by eliciting the client’s statements of desire, ability, reasons, and need for change. Then in phase 2, the focus shifts to strengthening commitment to change. Amrhein’s findings also converge with the commonsense precept that people tend to find their own verbalizations persuasive for guiding their behavior (Bem 1967, Hosford et al. 1995), and with more recent finding that stated implementation intentions predict behavioral follow-through, particularly when accompanied by a specific plan for carrying out the change (Gollwitzer 1999). These psycholinguistic data provided a missing piece in the emergent theory of MI, supporting the link between client in-session speech and posttreatment outcomes. We had been measuring the wrong statistic (intercept rather than slope) for the wrong metric (frequency instead of intensity) of the wrong dependent variable (generic change talk rather than commitment), and in the wrong portion of MI sessions (beginning rather than ending). The client’s starting level of motivation in an MI session was unrelated to outcome; it was commitment strength during the final minutes of the session that most strongly predicted behavior change (Amrhein et al. 2003).
Finally, research has addressed the question of optimal methods for helping clinicians to learn the intervention style of MI. Trainers are often asked to teach MI in periods varying from one hour to one day, and counselors sometimes attend such training in the hope of learning a few tricks to make clients do what they want them to do. MI is nothing of the sort. Rather, it is a complex clinical style for eliciting the client’s own values and motivations for change. It is far more about listening than telling, about evoking rather than instilling. MI communicates not, “I have what you need,” but instead, “You have what you need, and together we will find it.”

The most familiar vehicle for continuing professional education is the expert workshop, which in MI is often offered over the course of two full days. How effective are such workshops in increasing clinician proficiency in MI? This was the question addressed in an evaluation of a two-day workshop offered by Miller, with outcomes assessed not only by clinician self-report but also by practice samples obtained before and after training (Miller & Mount 2001). Participants submitted tape recordings of their counseling with actual clients prior to and several months after the workshop and interacted with a standard-patient actor to demonstrate their posttraining skill acquisition. After training, the clinicians showed modest albeit statistically significant increases in MI-congruent practice behavior, but not enough to make any difference in how their clients responded. Clients showed no change in levels of resistance or change talk after the clinicians were trained. On self-report, however, workshop participants reported confidence that they were now reasonably proficient in MI and were implementing it in practice. Such glowing self-reports of benefit from training are common (Rubel et al. 2000), but proved to be uncorrelated with actual increases in proficiency (Miller & Mount 2001).

In a subsequent trial of training methods, clinicians who wanted to learn MI were randomly assigned to receive or not to receive, in addition to the two-day workshop, one or both of two aids for learning: specific proficiency feedback from practice tapes, and six expert coaching consultations by telephone (Miller et al. 2005). A wait-list control group was given the MI book and training videotapes (Miller et al. 1998) and asked to improve their MI skills on their own, prior to attending the workshop. Based on Amrhein’s findings reported above, we also changed our training to a learning-to-learn format. We instructed trainees that they would not be skillful in MI by the end of the workshop, but that if we were successful they would know how to learn MI from their clients. Specific emphasis was placed on recognizing client speech events (change talk, commitment, resistance) that are relevant to behavioral outcomes, and using these as differential cues to shape successful practice.

As before, those receiving only the workshop showed modest gains in MI skills, and did not reach proficiency thresholds required for therapists in a clinical trial. Clinicians working on their own from MI tapes and book showed little improvement in skillfulness. Either or both of the training aids, however, significantly
improved post-workshop MI proficiency, and participants in these groups on average reached levels required for clinical trial certification.

SUMMARY

The evidence base for motivational interviewing is strong in the areas of addictive and health behaviors. Useful as a brief intervention in itself, MI also appears to improve outcomes when added to other treatment approaches. New research is clarifying the causal processes underlying the efficacy of motivational interviewing, and exploring optimal methods for helping practitioners to develop proficiency in this clinical method.

The Annual Review of Clinical Psychology is online at http://clinpsy.annualreviews.org

LITERATURE CITED

Gentilello LM, Rivara FP, Donovan DM,


CONTENTS

A HISTORY OF CLINICAL PSYCHOLOGY AS A PROFESSION IN AMERICA (AND A GLIMPSE AT ITS FUTURE), Ludy T. Benjamin, Jr. 1

STRUCTURAL EQUATION MODELING: STRENGTHS, LIMITATIONS, AND MISCONCEPTIONS, Andrew J. Tomarken and Niels G. Waller 31

CLINICAL JUDGMENT AND DECISION MAKING, Howard N. Garb 67

MOTIVATIONAL INTERVIEWING, Jennifer Hettema, Julie Steele, and William R. Miller 91

STATE OF THE SCIENCE ON PSYCHOSOCIAL INTERVENTIONS FOR ETHNIC MINORITIES, Jeanne Miranda, Guillermo Bernal, Anna Lau, Laura Kohn, Wei-Chin Hwang, and Teresa La Fromboise 113

CULTURAL DIFFERENCES IN ACCESS TO CARE, Lonnie R. Snowden and Ann-Marie Yamada 143

COGNITIVE VULNERABILITY TO EMOTIONAL DISORDERS, Andrew Mathews and Colin MacLeod 167

PANIC DISORDER, PHOBIAS, AND GENERALIZED ANXIETY DISORDER, Michelle G. Craske and Allison M. Waters 197

DISSOCIATIVE DISORDERS, John F. Kihlstrom 227

THE PSYCHOBIOLOGY OF DEPRESSION AND RESILIENCE TO STRESS: IMPLICATIONS FOR PREVENTION AND TREATMENT, Steven M. Southwick, Meena Vythilingam, and Dennis S. Charney 255

STRESS AND DEPRESSION, Constance Hammen 293

THE COGNITIVE NEUROSCIENCE OF SCHIZOPHRENIA, Deanna M. Barch 321

CATEGORICAL AND DIMENSIONAL MODELS OF PERSONALITY DISORDER, Timothy J. Trull and Christine A. Durrett 355

THE DEVELOPMENT OF PSYCHOPATHY, Donald R. Lynam and Lauren Gudonis 381

CHILD MALTREATMENT, Dante Cicchetti and Sheree L. Toth 409

PSYCHOLOGICAL TREATMENT OF EATING DISORDERS, G. Terence Wilson 439

GENDER IDENTITY DISORDER IN CHILDREN AND ADOLESCENTS, Kenneth J. Zucker 467
<table>
<thead>
<tr>
<th>CONTENTS</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>THE DEVELOPMENT OF ALCOHOL USE DISORDERS, <em>Kenneth J. Sher,</em></td>
<td>493</td>
</tr>
<tr>
<td><em>Emily R. Grekin,</em> and <em>Natalie A. Williams</em></td>
<td></td>
</tr>
<tr>
<td>DECISION MAKING IN MEDICINE AND HEALTH CARE, <em>Robert M. Kaplan</em></td>
<td>525</td>
</tr>
<tr>
<td>and <em>Dominick L. Frosch</em></td>
<td></td>
</tr>
<tr>
<td>PSYCHOLOGY, PSYCHOLOGISTS, AND PUBLIC POLICY,</td>
<td>557</td>
</tr>
<tr>
<td><em>Katherine M. McKnight,</em> <em>Lee Sechrest,</em> and <em>Patrick E. McKnight</em></td>
<td></td>
</tr>
<tr>
<td>COGNITIVE APPROACHES TO SCHIZOPHRENIA: THEORY AND THERAPY,</td>
<td>577</td>
</tr>
<tr>
<td><em>Aaron T. Beck</em> and <em>Neil A. Rector</em></td>
<td></td>
</tr>
<tr>
<td>STRESS AND HEALTH: PSYCHOLOGICAL, BEHAVIORAL, AND BIOLOGICAL</td>
<td>607</td>
</tr>
<tr>
<td>DETERMINANTS, <em>Neil Schneiderman,</em> <em>Gail Ironson,</em> and <em>Scott D. Siegel</em></td>
<td></td>
</tr>
<tr>
<td>POSITIVE PSYCHOLOGY IN CLINICAL PRACTICE, <em>Angela Lee Duckworth,</em></td>
<td>629</td>
</tr>
<tr>
<td><em>Tracy A. Steen,</em> and <em>Martin E. P. Seligman</em></td>
<td></td>
</tr>
<tr>
<td>INDEX</td>
<td>653</td>
</tr>
<tr>
<td>Subject Index</td>
<td></td>
</tr>
</tbody>
</table>