The Efficacy of Motivational Interviewing: A Meta-Analysis of Controlled Clinical Trials

Brian L. Burke
Fort Lewis College

Hal Arkowitz and Marisa Menchola
University of Arizona

A meta-analysis was conducted on controlled clinical trials investigating adaptations of motivational interviewing (AMIs), a promising approach to treating problem behaviors. AMIs were equivalent to other active treatments and yielded moderate effects (from .25 to .57) compared with no treatment and/or placebo for problems involving alcohol, drugs, and diet and exercise. Results did not support the efficacy of AMIs for smoking or HIV-risk behaviors. AMIs showed clinical impact, with 51% improvement rates, a 56% reduction in client drinking, and moderate effect sizes on social impact measures ($d = 0.47$). Potential moderators (comparative dose, AMI format, and problem area) were identified using both homogeneity analyses and exploratory multiple regression. Results are compared with other review results and suggestions for future research are offered.

Motivational interviewing is a relatively new and promising therapeutic approach that integrates the relationship-building principles of humanistic therapy (Rogers, 1951) with more active cognitive–behavioral strategies targeted to the client's stage of change (Prochaska, DiClemente, & Norcross, 1992). It has been defined as a client-centered yet directive method for enhancing intrinsic motivation to change by exploring and resolving client ambivalence (Miller & Rollnick, 2002). Since publication of the first edition of the motivational interviewing book (Miller & Rollnick, 1991), empirical research has advanced on approaches related to motivational interviewing for a variety of clinical problems. In the past decade, three exemplary studies that eliminated almost all threats to internal validity (Miller, Benefield, & Tonigan, 1993; Project MATCH, 1997; Stephens, Roffman, & Curtin, 2000) have provided strong support for the efficacy of these approaches in the areas of alcohol and drug addiction. The present article reviews this research domain meta-analytically, focusing on controlled clinical trials of individually delivered interventions that incorporated the four basic principles of motivational interviewing discussed in turn in the following text: (a) expressing empathy, (b) developing discrepancy, (c) rolling with resistance, and (d) supporting self-efficacy (Miller & Rollnick, 2002).

Although expressing empathy is fundamental to virtually all psychotherapies, in motivational interviewing it takes the specific form of reflective listening (or accurate empathy) as described by Carl Rogers (1951). Underlying this principle of empathy is a client-centered attitude of acceptance, wherein client ambivalence or reluctance to change is viewed as a normal part of the human experience rather than as pathology or pernicious defensiveness. Developing discrepancy, the second principle of motivational interviewing, is where it begins to depart from classic client-centered therapy. A key goal in motivational interviewing is to increase the importance of change from the client's perspective. This is accomplished using specific types of questions, along with selective reflections, that direct the client toward the discrepancy between his or her problem behavior and broader personal values.

Although motivational interviewing is intentionally directive, the therapist is careful not to explicitly advocate for change; it is the client who presents the reasons for change. Accordingly, when a client expresses resistance to change, it is a signal for the interviewer to respond differently. Resistance is conceptualized as an interpersonal variable, and the third basic principle of motivational interviewing is not to oppose the client's resistance actively but rather to accept and flow with it, again using reflective listening skills. Finally, a client's readiness for change is hypothesized to stem from two main factors: the importance of the change for the client (as discussed above) and the confidence the client has about successfully making the change. This confidence, often termed *self-efficacy*, is an essential element in motivation and a good predictor of treatment outcome (Bandura, 1997). The fourth guiding principle of motivational interviewing, therefore, is to enhance the client's confidence in his or her own capability to cope with obstacles and to succeed in changing.

In the research literature, the most widely used approach related to motivational interviewing has been one in which the client (often alcohol or drug addicted) is given feedback based on individual results from standardized assessment measures, such as the Drinker's Check-Up (Miller, Sovereign, & Krege, 1988) or a

modification of it. This feedback, which concerns the client's level of severity on the target symptom compared with norms, is delivered in a motivational interviewing "style," wherein possibilities for change are elicited from the client in a nonthreatening manner. Discussion of the problem may extend to one or more sessions that continue to embody the fundamental spirit and methods of motivational interviewing. We, along with Miller (W. R. Miller, personal communication, March 2001), consider this feedback-based approach to constitute an *adaptation* of motivational interviewing (AMI) because it is defined by the presence of the feedback component and not solely by the use of motivational interviewing per se. More broadly, we also apply the term *AMI* to interventions that incorporate additional nonmotivational interviewing techniques while retaining motivational interviewing principles as the core of treatment as well as to interventions that have been specifically adapted for use by nonspecialists (Rollnick, Heather, & Bell, 1992). To date, virtually all of the empirical studies in this area (and therefore in this review) have dealt with the efficacy of AMIs, and no studies have addressed the efficacy of motivational interviewing in its relatively pure form.

Three previous reviews of approaches related to motivational interviewing have been published. Noonan and Moyers (1997) reviewed the 11 clinical trials of AMIs available at that time (9 with problem drinkers and 2 with drug abusers) and concluded that 9 of these studies supported the efficacy of AMIs for addictive behaviors. Dunn, DeRoo, and Rivara (2001) performed a systematic review of 29 randomized trials of brief interventions that claimed to use the principles and techniques of motivational interviewing (or what we have called AMIs) to change behavior in four areas: (a) substance abuse, (b) smoking, (c) HIV-risk reduction, and (d) diet and exercise. Data on methodological features were tabled, as were calculations of effect sizes and their 95% confidence intervals, although the authors chose not to combine or compare data meta-analytically. The strongest evidence for efficacy was found in the alcohol and drug abuse areas, in which AMIs appeared to work well for problem drinkers and improved the rate of entry into and retention in intensive substance abuse treatment. AMI effects did not appear to diminish over time, and the effect sizes for AMIs as preludes to other treatments (e.g., inpatient care) were roughly equivalent to those for AMIs as stand-alone interventions.

More recently, Burke, Arkowitz, and Dunn (2002) qualitatively reviewed 26 studies that met their specified inclusion criteria. The authors concluded that the research supported the efficacy of AMIs for alcohol problems, drug addiction, hypertension, and bulimia as well as its efficacy for encouraging compliance in patients with diabetes. Mixed support was found for AMIs in the areas of reducing cigarette smoking, increasing physical activity, and enhancing dietary adherence in patients with hyperlipidemia. No support was found for AMIs in the reduction of HIV-risk behaviors (e.g., needle-sharing). In general, the AMIs reviewed were superior to no-treatment control groups and less credible alternative treatments and equal to active comparison treatments. After examining evidence regarding the mechanism of AMIs, Burke et al. (2002) reported that the research literature failed to shed light on how the treatment actually works. For instance, no direct support was found for the idea that AMIs exert their clinical effects by enhancing the client's motivation to change. In addition, the authors found virtually no data to indicate for whom these treat-

ments were optimal, as most clinical trials of AMIs that looked for aptitude by treatment interactions (Shoham & Rohrbaugh, 1995) were unable to find them. Thus, three previous reviews have examined the AMI literature using largely qualitative methodology, with no quantitative synthesis of this research domain to date. For the current article, we decided to review the AMI literature meta-analytically for three main reasons: (a) traditional methods of reviewing may suffer a considerable loss of power relative to meta-analytic methods, hence inflating the probability of Type II errors (Rosenthal, 1991); (b) meta-analytic reviews are likely to lead to summary statements of greater precision and objectivity (Kaplan, 1964); and (c) several controlled clinical trials of AMIs could be included that were not yet available for inclusion in any of the previous reviews.

To our knowledge, the current review is the first meta-analytic examination of the motivational interviewing literature. This meta-analysis is *multidimensional* (Westen & Morrison, 2001), providing a range of statistics bearing on outcome in addition to the usual effect sizes. Our review has five main objectives. First, we present the basic characteristics (e.g., problem types, settings, treatment lengths, comparison groups) of the controlled clinical trials of AMIs. Second, we evaluate the evidence for the efficacy of AMIs across clinical problem areas compared with control procedures (e.g., no treatment or weak alternatives) and with other active treatments. Third, these studies are examined for evidence of *sustained efficacy*, the ability of the AMIs to produce lasting symptomatic changes rather than an initial response solely (Westen & Morrison, 2001). Fourth, we explore evidence for the *clinical impact* of AMIs—the practical value or importance of an intervention to clients or to others with whom clients interact (Kazdin, 1999). Fifth, we focus on the identification of *moderator* variables, factors associated with variations in the outcome of controlled clinical trials of AMIs that may shed light on why different studies produced different results (Rosenthal & DiMatteo, 2001).

Method

Study Selection

For this review, we searched through the reference sections of all three prior reviews and the motivational interviewing website (www .motivationalinterview.org). We also conducted a database search (PsycINFO) using *motivational interviewing* as a key phrase, and finally, we sent out an electronic message to all members of the Motivational Interviewing Network of Trainers asking for any published or unpublished studies relevant to our purposes.

This meta-analysis follows the general guidelines commonly used in reviews of the efficacy of various psychotherapies (e.g., Kazdin, 1992). For this reason, studies had to satisfy the following criteria to be included in this review: (a) the intervention under study consisted primarily of implementing the motivational interviewing principles discussed above rather than principles of some other approach (such as cognitive–behavioral therapy), (b) the intervention was delivered on an individual (i.e., not group) and face-to-face (i.e., not telephone) basis, and (c) the study design met our criteria for a controlled clinical trial. In our definition, a controlled clinical trial must use the following: (a) random assignment to groups or an alternative way of equating groups of clients before treatment (e.g., sequential assignment), (b) at least one comparison group, and (c) adequate measurement targeting pertinent problem areas. Although the controlled clinical trial has recently come under criticism (e.g., Borkovec & Caston-

guay, 1998), it remains the gold standard for evaluating treatment outcome (Stanton & Shadish, 1997).

Statistical Analyses

Our general data analytic approach for this review is outlined in Table 1. The specific strategies and procedures used in implementing this approach are discussed in detail in the following text.

What Kind of Controlled Trials Have Been Done With AMIs?

The first goal of this meta-analysis was to characterize the types of questions that have been investigated in controlled trials of AMIs. Accordingly, we coded descriptive characteristics relating to important methodological and substantive features of each study. Substantive features included problem area as well as treatment setting, format, and dose, whereas methodological features included sample size, study design (i.e., type of control and/or comparison groups used), dependent measures used, and follow-up lengths and rates.

What Is the Comparative Efficacy of AMIs?

For each study reporting sufficient information, we calculated effect sizes and confidence intervals for the main behavioral and health outcomes at all reported follow-up times. When necessary, authors were contacted for group means and standard deviations not reported in the original article (Gentilello et al., 1999; Juárez, 2001; Martino, Carroll, O'Malley, & Rounsaville, 2000). All author inquiries yielded useful data for effect size calculations. For each follow-up interval of each treatment comparison involving an AMI, a unit-free effect size, g, was calculated by subtracting the control group mean from the experimental (AMI) group mean and dividing the result by the pooled standard deviation according to the following formula (Hedges & Olkin, 1985, pp. 78–79):

$$\text{Effect size} = g = (Y_E - Y_C)/s$$

$$\text{and } s = \sqrt{[(n_E - 1)(s_E)^2 + (n_C - 1)(s_C)^2]/[n_E + n_C - 2]} \quad (1)$$

where Y_E and Y_C are the experimental and control group means posttreatment, s is the pooled standard deviation, s_E and s_C are the experimental and control group standard deviations, and n_E and n_C are the experimental and control group sample sizes. In all cases, an unbiased estimate of the

population effect size was then obtained by correcting for the bias in g (Hedges & Olkin, 1985, p. 81). When means were not available, the effect size was estimated directly from significance tests (t, F, or chi-square) according to the requisite procedures (for more details, see Rosenthal, 1991, pp. 18–20).

For all effect sizes, 95% confidence intervals were derived from their variance, which was estimated according to the following formula (Hedges & Olkin, 1985, p. 86):

$$\sigma^2(d) = [(n_E + n_C)/n_E n_C] + [d^2/2(n_E + n_C)] \quad (2)$$

where n_E and n_C are the experimental and control group sample sizes respectively.

How Do AMIs Compare With No-Treatment or Placebo Groups and With Other Active Treatments Across Different Problem Areas?

Prior to generating combined effect size estimates, effect sizes were grouped into theoretically meaningful subcategories according to the following two variables: (a) clinical problem area (alcohol, smoking, drug addiction, HIV-risk behaviors, or diet and exercise) and (b) design type (no-treatment and/or placebo control or active treatment comparison group). A group of independent effect size estimates was then generated within each category (e.g., AMI vs. no-treatment and/or placebo control for alcohol problems). For studies using multiple follow-up points, the first posttreatment effect size was selected. For studies using multiple outcome measures, the effect size associated with the best *target measure* was selected a priori according to psychometric properties and common usage (e.g., standard drinks/week as a measure of alcohol consumption). When possible, combined effect sizes were also generated separately for secondary target measures, such as peak blood alcohol concentration for alcohol problems.

The combined effect size (d_c), was computed by weighting each individual effect size according to the inverse of its variance. In this way, each study contributed to the combined estimate according to the precision of its own effect size estimates (i.e., studies with larger sample sizes contributed more heavily to the combined effect size). For each combined effect size, 95% confidence intervals were derived from the variance of d_c (Hedges & Olkin, 1985, p. 113).

Table 1
General Data Analytic Approach for Review

Key question	Data analytic strategy
1. What kind of controlled trials have been done with adaptations of motivational interviewing (AMIs)?	Code descriptive characteristics of all 30 studies
2. What is the comparative efficacy of AMIs?	Compute individual effect sizes for each study
(a) How do AMIs compare with no-treatment or placebo groups across different problem areas?	Compute combined effect sizes for AMIs versus no-treatment/placebo comparison groups by problem area;
(b) How do AMIs compare with other active treatment across different problem areas?	compute combined effect sizes for AMIs versus active treatment comparison groups by problem area
3. What is the sustained efficacy of AMIs?	Compare posttreatment and follow-up effect sizes of AMIs
4. What is the clinical impact of AMIs?	For AMI studies, whenever available, combine percentage improved data; compute and combine within-AMI effect sizes; combine drinking frequency data from alcohol studies; compute combined effect sizes for social impact measures
5. What factors might account for any observed differences in effect sizes across these studies?	If Q tests indicate significant heterogeneity, subdivide effect size groupings further using categorical moderators; code and analyze potential moderators by multiple regression

What Is the Sustained Efficacy of AMIs?

To examine whether the efficacy of AMI interventions was sustained over time, we conducted the most rigorous test possible by examining only those studies that generated effect sizes at both posttreatment and at least one other follow-up period. In this way, the difference between posttreatment and follow-up results is not confounded with any other study features (Stanton & Shadish, 1997). For this subset of studies, Gleser and Olkin's (1994) procedure for stochastically dependent data was used. This analysis yielded a chi-square statistic that could be tested for the significance of the difference between posttreatment and long-term effects of the AMI interventions.

What Is the Clinical Impact of AMIs?

Several AMI studies provided data on the percentage of clients who improved or who abstained from alcohol and drugs following treatment. These data were combined meta-analytically to yield percentage improved estimates for alcohol and drug addiction that ranged from 1 month to 4 years posttreatment. Percentage improved estimates provide another measure of sustained treatment efficacy (Westen & Morrison, 2001) as well as an indicator of clinical impact (Kazdin, 1999). One caveat to note is that *improved* was defined somewhat differently by various authors, with definitions that ranged from "abstinence or asymptomatic moderate drinking" in Brown and Miller (1993) to "at least a 50% reduction in drinking from baseline" in Miller et al. (1988). Even *abstinence* was not uniformly defined across studies, as it was used to refer to zero substance use at the time of follow-up (as verified by a significant other) in Bien, Miller, and Boroughs (1993) or, more stringently, to zero substance use in the 1 (Wertz, 1994), 3 (Stephens et al., 2000), or 6 (Schneider, Casey, & Kohn, 2000) months before the follow-up point.

For all studies providing the necessary data*,* within-treatment (pre–posttreatment) effect sizes were computed for groups receiving AMIs as stand-alone interventions. These within-AMI effect sizes were then combined by weighting each one according to its sample size. Although widely reported, within-treatment effect sizes are not so meaningful because these estimates confound real treatment effects with spurious effects such as regression to the mean and placebo factors, hence undermining the causal inferences that are the virtue of experimental designs (Westen & Morrison, 2001). Despite this serious caveat, within-AMI effect sizes are presented here as another indicator of clinical impact, providing an estimated upper limit of the effects of AMIs before alternative explanations for these effects have been ruled out.

Several studies provided data on the drinking frequency of clients (in standard ethanol content [SEC] per week) before and after stand-alone AMI treatments. These data were combined meta-analytically to yield estimates for mean alcohol consumption pre- and post-AMI treatment, thereby providing an additional indicator of clinical impact.

For alcohol and drug abuse studies, combined effect sizes of AMIs were estimated separately for social impact measures (Kazdin, 1992). These are outcome measures that tap into broad life problems related to the target symptom. Social impact measures used in these studies were the Rutgers Alcohol Problem Index (White & Labouvie, 1989), five items from the Health Behavior Questionnaire (Jessor, Donovan, & Costa, 1989), and scales tailored to measure opiate-related problems (adapted by Saunders, Wilkinson, & Phillips, 1995, from Skinner, 1982) and marijuana-related problems (Stephens et al., 2000). These measures included variables such as substance-related work or academic impairment, physical symptoms (e.g., memory loss, injuries), legal problems (e.g., driving under the influence), and social impairments, serving as yet another indicator of the clinical impact of the AMI interventions.

What Factors Might Account for Any Observed Differences in Effect Sizes Across These AMI Studies?

Potential moderators.   In a meta-analysis, the potential moderators to be investigated are necessarily constrained by what the original authors chose to measure and include in their research reports. We first generated a list of eight key variables that were reported in at least 90% of reviewed studies: problem area, treatment setting, client diagnoses or severity information, format of the AMI treatment, treatment doses, methodology and design features, follow-up intervals, and therapist training. From this list, we selected six variables that might, on theoretical grounds, be expected to moderate the effects of AMI treatment:

1. Clinical problem area, either alcohol and/or drug addiction or diet and exercise. Note that, because of low study numbers, some problem types were excluded from the moderator analyses (i.e., smoking, HIV risk, treatment adherence, and eating disorders). Rationale: As may be the case in psychotherapy, a particular technique or approach (e.g., AMI) may be a good fit for certain types of problems but not for others.

2. Severity of alcohol or drug addiction, which was coded by two independent raters as follows (for substance abuse studies only): (a) mild, if the sample at intake consumed less than 20 standard drinks/week or was diagnosed with substance abuse according to *Diagnostic and Statistical Manual of Mental Disorders* (*DSM*; 3rd ed., rev. [*DSM–III–R*]; American Psychiatric Association, 1987) or *DSM–IV* (4th ed.; American Psychiatric Association, 1994); (b) mixed/moderate, if the sample at intake consumed between 20 and 60 standard drinks/week or was mixed in terms of substance abuse and substance dependence diagnoses (*DSM–III–R* or *DSM–IV*); or (c) severe, if the sample at intake consumed more than 60 standard drinks/week or was diagnosed with substance dependence (*DSM–III–R* or *DSM–IV*). The Spearman–Brown formula (Rosenthal, 1991, p. 52) was used to estimate the effective interrater reliability (R) of these severity ratings, which was reasonably good (R = .83). Rationale: Studies in other areas (e.g., Elkin, Gibbons, Shea, & Sotsky, 1995) have found differential treatment outcomes for clients with high and low problem severity, and we wanted to examine this possibility with AMIs as well.

3. Format of the AMI, either a stand-alone intervention or a prelude and/or adjunct to further clinical services. Rationale: Although both of these formats fit our definition of an AMI, these two uses are quite different procedurally and each answers a different question. Using an AMI as a stand-alone treatment allows one to examine to what degree it is efficacious as an intervention in itself, whereas use of an AMI as a prelude and/or adjunct to further services allows one to examine the degree to which it may be a catalyst that enhances the efficacy of another treatment approach.

4. Comparative dose of the AMI treatment, measured in total minutes of AMI (minutes per session times number of sessions) minus total minutes of comparison (or control) treatment. Rationale: The psychotherapy literature has shown that a treatment's dose can indeed moderate its efficacy (Howard, Kopta, Krause, & Orlinsky, 1986; Shadish, Navarro, Matt, & Phillips, 2000).

5. Methodological quality of the study, with two independent raters judging the quality of each study on 12 specific dimensions (presented in Table 2). Some dimensions (e.g., group allocation, follow-up rate and length) counted for more than 1 point, which resulted in a maximum possible methodological quality score

Table 2
Methodological Quality Score Coding System

Methodological feature/dimension	Points awarded
Group allocation	4 = true randomization 3 = within-subject counterbalanced 2 = case control/matching 1 = quasiexperimental design; arbitrary/sequential assignment 0 = violated randomization or nonequivalent groups
Quality control	1 = treatment standardized by manual, specific training, and so forth 0 = no standardization specified
Follow-up rate	2 = 85%–100% of follow-ups completed 1 = 70%–84.9% of follow-ups completed 0 = < 70% of follow-ups completed or follow-up length < 3 months
Follow-up length	2 = 12 months or longer 1 = 6.0–11.9 months 0 = < 6 months or unspecified
Contact	1 = personal or telephone contact for > 70% of completed follow-ups 0 = questionnaire, unspecified, or < 70% of follow-ups contacted in person or by phone
Collaterals	1 = collaterals (e.g., the client's significant others) interviewed in > 50% of cases 0 = No collateral verification in most cases, or unspecified
Objective	1 = objective verification (records, serum, breath, etc.) in > 50% of cases 0 = no objective verification in most cases, or unspecified
Dropouts	1 = treatment dropouts included in at least some outcome data (e.g., intent to treat analysis; compared on dependent variable; etc.) 0 = treatment dropouts not discussed or not accounted for (e.g., excluded noncompleters from all analyses)
Attrition	1 = cases lost to follow-up enumerated and considered in outcome reporting (e.g., counted as failures; compared with nonattrition cases on prior characteristics) 0 = lost cases not enumerated or merely enumerated but not considered in outcome
Independent	1 = follow-up done by independent interviewer 0 = follow-up nonblind, unspecified, or questionnaire only
Analyses	1 = acceptable statistical analyses of group differences 0 = no statistical analyses, inappropriate analyses, or unspecified
Multisite	1 = parallel replications at two or more sites with separate research teams 0 = single site or comparison of sites offering different treatments

Note. This methodological quality score coding system was developed by Miller et al. (1995).

(MQS) of 17 (see Miller et al., 1995, for more details on this rating system). The Spearman–Brown formula (Rosenthal, 1991, p. 52) was used to estimate the effective interrater reliability of the total MQS for each study, which was reasonably high (R = .86). Rationale: Although many previous meta-analyses have failed to find a strong relationship between methodological quality and study outcome (e.g., Glass, McGaw, & Smith, 1981), several meta-analyses in the substance abuse literature have suggested that there may in fact be an inverse relationship between these two variables (e.g., Garrett, 1985; Stanton & Shadish, 1997).

6. Follow-up point, measured in weeks posttreatment. Rationale: One of the most critical questions about the efficacy of any treatment is whether there is a difference between its short- and long-term outcomes.

Homogeneity analyses. Several homogeneity analyses were performed as described below, each yielding a Q statistic with an asymptotic chi-square distribution (Hedges & Olkin, 1985, p. 123). A nonsignificant Q indicates that the variance in the distribution of effect sizes may be attributed to sampling error, whereas a significant Q indicates that there is some consequential variation or heterogeneity across studies meriting further exploration (Hunter, Schmidt, & Jackson, 1982). The result of each homogeneity analysis was therefore used to decide whether to further subdivide that particular group of effect sizes according to moderator variables (e.g., dose of AMI); this procedure was repeated until all subgroups were statistically homogeneous.

Multiple regression. In addition to subdividing heterogeneous effect size groupings according to categorical moderators as described previously, potential moderator variables were analyzed by means of weighted multiple regression. Because some moderators (i.e., severity) were specific to substance abuse studies (alcohol or drugs) and others referred to all studies (i.e., problem area), separate multiple regressions were performed for (a) substance abuse studies only and (b) all studies combined. For both regression analyses, we implemented an exploratory approach (forward selection followed by backward elimination) using a fixed effects SPSS macro (Lipsey & Wilson, 2001, pp. 138, 216), yielding a final predictive model in each case.

Reducing and Estimating Bias

Dealing With Subjectivity

As in all meta-analyses (and reviews in general), subjective judgment was involved, and there were decisions to be made at various points throughout the process. To minimize allegiance effects in this article, we used the approach of Westen and Morrison (2001). When two procedures seemed equally reasonable and statistically sound, we selected the one that would minimize the effect size estimation of the AMI interventions. For example, we used the posttreatment pooled standard deviation (whenever available) as the denominator in calculating effect sizes, which produced smaller estimates than if we had used the pretreatment pooled standard deviation instead. Thus, the effect sizes and confidence intervals reported in this article may be considered conservative tests of the hypothesis that AMIs are efficacious.

Accounting for Attrition

Another potential source of bias in meta-analyses is attrition, which merits scrutiny because it can present a significant threat to the internal validity of a controlled experiment (Liddle & Dakof, 1995). We wanted to see whether taking attrition into account would alter the conclusions we reached. Accordingly, we computed an additional effect size, total attrition *d* (TA *d*), for AMI studies compared with no-treatment and/or placebo controls whenever sufficient data were available. This was calculated by counting all participants who were lost to follow-up as failures and by assigning each of them the posttreatment mean of the control group as an individual outcome score. The resulting effect size (TA *d*) is equivalent to the classic intent-to-treat analysis, in which all participants assigned to conditions are included (Stanton & Shadish, 1997) and therefore provides an estimated lower limit for the efficacy of AMIs in these studies.

File Drawer Problem

Statisticians and behavioral researchers have long suspected that the studies published in journals are a biased sample of the studies that are actually carried out (Rosenthal, 1991). To establish reasonable boundaries on this file drawer problem, we calculated the number of unavailable (filed or future) studies averaging null results that would reduce our findings to a nonsignificant level (see Rosenthal, 1991, p. 104). This result was then used to estimate whether publication bias was a threat to the overall conclusions of this review.

Investigator Allegiance Effects

A final source of bias evaluated in this review was investigator allegiance (Luborsky et al., 1999). It is widely known that treatments tend to yield better empirical results when the researchers have a theoretical or personal investment in the approach. Some of the studies in this review were conducted in the clinic of W. R. Miller, the founder of motivational interviewing. To estimate this bias, we compared the average effect size of studies conducted in W. R. Miller's substance abuse clinic (alcohol studies only) with those conducted elsewhere, using an SPSS macro for weighted analysis of variance (ANOVA; Lipsey & Wilson, 2001, pp. 135, 212), the meta-analytic analog to a one-way ANOVA or *t* test.

Results

What Kind of Controlled Trials Have Been Done With AMIs?

Thirty controlled clinical trials involving AMIs met inclusion criteria for this meta-analysis. The basic characteristics of these studies are shown (grouped by clinical problem area) in Table 3. There were 15 studies investigating AMIs for alcohol problems, 2 for smoking cessation, 5 for drug addiction, 2 for HIV-risk behaviors, 4 for diet and exercise problems, and 1 each for treatment adherence and eating disorders. The most common settings for these clinical trials were substance abuse clinics (11) and hospitals (9), with some studies conducted in general medical practices (3) or on college campuses (3). Sample sizes of the studies ranged from 22 to 952, with a mean of 206 participants. Two different design types were used in these studies: In 26 studies, AMIs were compared with control groups, whereas in 9 studies, AMIs were compared with other bona fide, active treatments (5 studies used both design types). Control groups included 11 true no-treatment groups, 9 placebos (e.g., 5-min interview, informational letter), and 6 standard treatment groups (e.g., brief dietary advice). Active treatment comparison groups included 7 cognitive–behavioral in-

terventions (e.g., relapse prevention, risk reduction), 2 12-step facilitation groups, 2 confrontative feedback interventions, and 1 client-centered counseling condition.

The dependent variable was virtually always the target symptom (e.g., drinking frequency) and, in 8 studies, also included social impact measures tapping into symptom-related problems (e.g., days of work lost because of drinking). The AMI treatments were used either as stand-alone interventions (16 studies) or as preludes to further clinical services (14 studies), and the overall dose of the AMIs ranged from 15 to 240 min with a mean of 99 total treatment min. The *comparative* dose of the AMIs (minutes of AMI minus minutes of comparison treatment) ranged from −480 to 270 with a mean of 22 comparative min. The experimental rigor of the clinical trials was highly variable, as MQS ranged from 6 to 17 with a mean of about 10. Lengths and rates of follow-up were also heterogeneous in these studies: Follow-up length ranged from 4 weeks to 4 years, with a mean of 18 weeks, whereas rates of follow-up ranged from 26% to 98%, with a mean of 74%.

What Is the Comparative Efficacy of AMI?

A principal goal of this review was to explore the comparative efficacy of AMIs compared with control groups and other active treatments. Individual effect sizes for the AMI clinical trials are shown in Table 4. In this review, effect size means that a person receiving the AMI treatment improved by an average of *d* standard deviations on that particular measure (from intake to posttreatment) relative to someone in the control group. Of the 30 studies, only 11 produced at least one statistically significant (nonzero) effect size in favor of the AMI under investigation.

How Do AMIs Compare With No-Treatment or Placebo Groups and With Other Active Treatments Across Different Problem Areas?

Table 5 shows the combined effect sizes (with 95% confidence intervals) of AMIs across five clinical problem areas. Two fundamental summary results emerge from a close inspection of the table. First, while AMIs have shown significant (nonzero) effects compared with no-treatment/placebo controls, they have not produced any significant effects relative to other bona fide treatments. For alcohol and drug abuse studies, AMIs have yielded a combined effect size near zero (0.02) in the seven studies that compared them with other active treatments, although the AMIs were shorter than the alternative treatments by an average of 180 min (three or four sessions).

Second, the combined effect sizes of AMIs compared with no-treatment or placebo controls have been quite variable across the five problem areas. For illustrative purposes, these effect sizes are shown graphically in Figure 1. According to Cohen's (1988) classification system, AMIs have yielded medium effects in the areas of drug addiction (0.56) and diet and exercise (0.53). In the alcohol area, AMI effects have been small to medium (0.25 to 0.53), depending on the target measure (SEC or blood alcohol concentration). Conversely, AMIs have not shown any significant effects in the areas of smoking cessation or HIV-risk behaviors to date.

Table 3
Basic Characteristics of Controlled Clinical Trials Involving Adaptations of Motivational Interviewing (AMIs)

Study	Setting	N	Problem area	Severity	Design type	Measurement class	AMI format	Study rating (MQS)[a]	Dose of AMI (total min)	Longest follow-up interval (% completion)
Aubrey (1998)[b]	SAC	80	Alcohol	S2	D1	M1	P	7	45	3 months (49)
Bien, Miller, & Boroughs (1993)[b]	VA SAC	32	Alcohol	S3	D1	M1	P	13	60	6 months (72)
Borsari & Carey (2000)	College campus	60	Alcohol	S1	D1	M1, M2	S	9	60	6 weeks (98)
Brown & Miller (1993)[b]	Hospital	28	Alcohol	S3	D1	M1	P	10	100	3 months (89)
Gentilello et al. (1999)	Trauma center	762	Alcohol	S2	D1	M1	S	12	30	3 years
Handmaker, Miller, & Manicke (1999)[b]	Prenatal clinic	42	Alcohol	S1	D1	M1	S	9	60	2 months (81)
Heather et al. (1996)	Hospital	174	Alcohol	S2	D1, D2	M1	S	9	35	6 months (70)
Juárez (2001)[b]	College campus	122	Alcohol	S1	D1	M1, M2	S	10	70/50/50[d]	2 months (73)
Marlatt et al. (1998)	College campus	348	Alcohol	S1	D1	M1, M2	S	13	60	4 years (83)
Miller, Benefield, & Tonigan (1993)[b]	SAC	42	Alcohol	S2	D1, D2	M1	S	14	120	12 months (83)
Miller, Sovereign, & Krege (1988)[b]	SAC	42	Alcohol	S2	D1	M1	S	8	120	18 months (76)
Monti et al. (1999)	Hospital ER	94	Alcohol	S1	D1	M2	S	13	35	6 months (89)
Project MATCH Research Group (1997, 1998)	5 SACs	952/774[c]	Alcohol		D2	M1, M2	S/A	17	240	3 years (85)
Sellman et al. (2001)	SAC	125	Alcohol	S3	D1, D2	M1	S	12	240	6 months (98)
Wertz (1994)	Inpatient VA	42	Alcohol		D1	M1	P	6	60	1 month (52)
Butler et al. (1999)	GMP	536	Smoking cessation		D1	M1	S	10	15	6 months (78)
Colby et al. (1998)	Hospital	40	Smoking cessation		D1	M1	S	10	30	3 months (95)
Booth et al. (1998)	SAC	192	Drug addiction		D2	Treatment entry	P	10	120	3 months
Martino et al. (2000)	Hospital	23	Drug addiction	S2	D1	M1	P	6	50	12 weeks (26)
Saunders, Wilkinson, & Phillips (1995)	SAC	122	Drug addiction	S3	D1	M1, M2	P	9	60	6 months (60)
Schneider, Casey, & Kohn (2000)	EAP sites	89	Drug addiction		D2	M1, M2	S/P	9	120	9 months (71)
Stephens, Roffman, & Curtin (2000)	SAC	291	Drug addiction	S3	D1, D2	M1, M2	S	14	180	16 months (89)
Baker et al. (1993)	SAC	95	HIV-risk behaviors		D1, D2	M1	A	12	75	6 months (84)
Baker et al. (1994)	SAC	200	HIV-risk behaviors		D1	M1	S	10	30	6 months (44)
Swanson, Pantalon, & Cohen (1999)	Hospital	121	Treatment adherence	N	D1	% attending FAA	A	9	60	
Harland et al. (1999)	GMP	523	Diet & exercise	N	D1	M1	A	11	240/40[e]	1 year (85)
Mhurchu, Margetts, & Speller (1998)	Hospital dietary clinic	121	Diet & exercise	N	D1	M1	S	10	150	3 months (80)
Smith et al. (1997)	Outpatient clinic	22	Diet & exercise		D1	M1	A	10	150	4 months (73)
Woollard et al. (1995)	GMP	166	Diet & exercise	N	D1	M1	A	8	240	18 weeks (80)
Treasure et al. (1999)	Eating disorders clinic	125	Eating disorder		D2	M1	P	8	200	4 weeks (54)

Note. A = adjunct to standard treatment; D1 = no-treatment or placebo control group; D2 = active treatment comparison group; EAP = employee assistance program; ER = emergency room; FAA = first aftercare appointment; GMP = general medical practice; M1 = target symptom; M2 = social impact; P = prelude to further treatment; S = stand-alone treatment; S1 = mild; S2 = mixed/moderate; S3 = severe; SAC = substance abuse clinic; VA = Veterans Administration hospital.
[a] MQS = methodological quality score (range 0–17). [b] This study was conducted in W. R. Miller's clinic. [c] This study involved two separate samples (outpatient and aftercare). [d] This study used three separate AMI groups. [e] This study used two separate AMI groups.

Table 4
Effect Sizes, Percentage Improved, and Drinking Frequency From Controlled Clinical Trials of Adaptations of Motivational Interviewing (AMIs)

Study	Problem area	Treatment group: comparison group	Outcome measure	Time point (weeks)	d (95% CI)	TA d (95% CI)	Within AMI d	% improved (AMI group)	Drinks/week (SEC) pre- to post-AMI
Brown & Miller (1993)	Alcohol	AMI + ST: ST	SEC	13	**0.99 (0.16, 1.82)**	**0.92 (0.09, 1.74)**			
			BAC	13	0.67 (−0.13, 1.48)	0.63 (−0.18, 1.43)		64	
Bien, Miller, & Boroughs (1993)	Alcohol	AMI + ST: PL + ST	SEC	13	0.70 (−0.09, 1.49)	**0.61 (−0.18, 1.40)**			
				26	0.33 (−0.49, 1.15)	0.25 (−0.57, 1.07)		64	
			BAC	13	0.74 (−0.05, 1.54)	**0.65 (−0.14, 1.44)**			
				26	0.30 (−0.52, 1.12)	0.23 (−0.60, 1.05)			
Borsari & Carey (2000)	Alcohol	AMI: NT	SEC	6	**0.57 (0.05, 1.09)**	**0.57 (0.05, 1.09)**	0.75		17.6 to 11.4
			RAPI	6	−0.07 (−0.58, 0.44)	−0.07 (−0.58, 0.44)			
			Drinking frequency	26	**0.23ᵃ (0.01, 0.46)**	—			
Marlatt et al. (1998)	Alcohol	AMI: NT		208	0.07 (−0.16, 0.30)	0.06 (−0.17, 0.29)			
			Drinking consequences	52	**0.34 (0.11, 0.56)**	**0.28 (0.05, 0.50)**	0.20	67	
				208	**0.28 (0.04, 0.51)**	**0.23 (0.00, 0.46)**			
Gentilello et al. (1999)	Alcohol	AMI: NT	SEC	52	0.09 (−0.10, 0.29)	0.05 (−0.15, 0.24)		48% reduction in new injuries (vs. NT)	40.1 to 18.3
Miller, Sovereign, & Krege (1988)	Alcohol	AMI: NT	SEC	6	−0.03ᵇ (−0.80, 0.74)	N/A		40	
Miller, Benefield, & Tonigan (1993)	Alcohol	AMI: NT	SEC	6	0.64 (−0.11, 1.40)	0.64 (−0.11, 1.40)			
			BAC	6	0.55 (−0.21, 1.30)	0.55 (−0.21, 1.30)	0.68		49 to 12
				6	0.26 (−0.49, 1.00)				
			SEC	52	0.45 (−0.30, 1.20)				
		AMI: CFB		6	0.10 (−0.64, 0.84)				
			BAC	52	0.06 (−0.68, 0.80)				
Monti et al. (1999)	Alcohol	AMI: ST	ARP	52	0.43 (−0.01, 0.86)	0.38 (−0.05, 0.82)			
Handmaker, Miller, & Manicke (1999)	Alcohol	AMI: PL	SEC	9	0.03ᵃ (−0.64, 0.71)	—			
			BAC	9	**0.70ᵃ (0.01, 1.40)**	—	0.46		
Heather et al. (1996)	Alcohol	AMI: NT	SEC	26	0.16 (−.29, .60)	**0.12 (−.33, .56)**	0.80		49.5 to 27.6
		AMI: SBC	SEC	26	0.35 (−.07, .76)				

Table 4 (*continued*)

Study	Problem area	Treatment group: comparison group	Outcome measure	Time point (weeks)	d (95% CI)	TA d (95% CI)	Within AMI d	% improved (AMI group)	Drinks/week (SEC) pre-AMI to post-AMI
Project MATCH Research Group (1997, 1998)[c]	Alcohol	AMI: CBT	GGT	65	0.05 (−0.14, 0.24)			44	
			Drinking consequences	65	−0.01 (−0.20, 0.19)				
		AMI: TSF	GGT	65	−0.08 (−0.26, 0.11)				
			Drinking consequences	65	−0.18 (−0.37, 0.01)				
Project MATCH Research Group (1997, 1998)[d]	Alcohol	AMI + ST: CBT + ST	GGT	65	0.24 (−0.04, 0.52)			51	
			Drinking consequences	65	0.09 (−0.12, 0.31)				
		AMI + ST: TSF + ST	GGT	65	0.21 (−0.08, 0.50)				
			Drinking consequences	65	0.16 (−0.05, 0.38)				
Sellman et al. (2001)	Alcohol	AMI: BF	Heavy drinking 1 + times (%)	26	0.33[a] (−0.11, 0.77)	**0.33[a] (−0.11, 0.77)**			
		AMI: BF + CC	Heavy drinking 1 + times (%)	26	0.34[a] (−0.10, 0.78)				
Aubrey (1998)	Alcohol	AMI: NT	Total alcohol use days	13	**0.93 (0.24, 1.61)**	0.61 (−0.06, 1.28)			
Wertz (1994)	Alcohol	AMI: NT	SEC	4	—	—		55	
		AMI[e]: NT	SEC	9	0.00 (−0.66, 0.66)	**0.00 (−0.66, 0.66)**			
			BAC	9	0.17 (−0.50, 0.83)	0.10 (−0.56, 0.77)			
			RAPI	9	0.17 (−0.49, 0.84)	0.11 (−0.55, 0.77)			
Juárez (2001)	Alcohol	AMI[f]: NT	SEC	9	0.46 (−0.17, 1.10)	0.32 (−0.31, 0.95)	0.76		12.7 to 5.2
			BAC	9	0.49 (−0.15, 1.13)	0.34 (−0.29, 0.97)			
			RAPI	9	−0.11 (−0.74, 0.52)	−0.08 (−0.71, 0.55)			
		AMI[g]: NT	SEC	9	0.44 (−0.23, 1.11)	**0.29 (−0.38, 0.95)**			
			BAC	9	0.12 (−0.54, 0.79)	0.08 (−0.58, 0.74)			
			RAPI	9	0.14 (−0.52, 0.80)	0.09 (−0.57, 0.75)			
Colby et al. (1998)	Smoking	AMI: PL	FTQ	13	0.24 (−0.40, 0.88)	**0.22 (−0.42, 0.85)**	0.18		
			Cigarettes/day	13	−0.04 (−0.67, 0.60)	−0.03 (−0.67, 0.61)			

(*table continues*)

Table 4 (*continued*)

Study	Problem area	Treatment group: comparison group	Outcome measure	Time point (weeks)	d (95% CI)	TA d (95% CI)	Within AMI d	% improved (AMI group)	Drinks/week (SEC) pre- to post-AMI
Butler et al. (1999)	Smoking	AMI[h]: PL	Abstinence past month (%)	26	**0.10[a] (−0.07, 0.27)**	**0.10[a] (−0.07, 0.27)**			
			Abstinence past day (%)	26	**0.22[a] (0.05, 0.39)**	**0.22[a] (0.05, 0.39)**			
Schneider, Casey, & Kohn (2000)	Drugs	AMI: CFB	Alcohol ASI	13	0.24 (−0.21, 0.69)		0.86	21	34.6 to 5.6
				39	0.43 (−0.08, 0.93)				
			Drug ASI	13	0.25 (−0.20, 0.70)				
				39	0 (−0.50, 0.50)				
			Substance-related problems	13	−0.02[a] (−0.47, 0.43)				
				39	−0.10[a] (−0.60, 0.40)				
Martino et al. (2000)	Drugs	AMI + ST: PL + ST	Frequency of substance use	12	0.41 (−0.42, 1.24)	—			
			Treatment days attended	12	0.46 (−0.37, 1.30)	—			
Booth et al. (1998)	Drugs	AMI: RR	Treatment entry (%)	13	−0.06[a] (−0.34, 0.22)				
Saunders, Wilkinson, & Phillips (1995)	Drugs	AMI: PL	Opiate dependence	26	0.06[a] (−0.40, 0.52)	—			
			Opiate-related problems	26	**0.48[a] (0.02, 0.95)**	—			
Stephens, Roffman, & Curtin (2000)	Drugs	AMI: RP Group	Frequency of marijuana use	18	−0.11 (−0.42, 0.19)		2.60	37	
				70	−0.06 (−0.35, 0.23)				
			Marijuana-dependent problems	18	0.06 (−0.24, 0.36)				
				70	−0.10 (−0.40, 0.19)				
		AMI: NT	Frequency of marijuana use	18	**0.84 (0.51, 1.17)**	**0.72 (0.39, 1.05)**			
			Marijuana-dependent problems	18	**1.12 (0.78, 1.46)**	**0.95 (0.62, 1.29)**			

Table 4 (*continued*)

Study	Problem area	Treatment group: comparison group	Outcome measure	Time point (weeks)	d (95% CI)	TA d (95% CI)	Within AMI d	% improved (AMI group)	Drinks/week (SEC) pre- to post-AMI
Baker et al. (1993)	HIV-risk behavior	AMI + ST + RP: AMI + ST	HIV risk-taking behavior	26	−0.15 (−0.69, 0.39)	−0.15 (−0.67, 0.42)			
		AMI + ST: PL + ST	HIV risk-taking behavior	26	−0.15 (−0.70, 0.39)	−0.12 (−0.67, 0.42)			
Baker et al. (1994)	HIV-risk behavior	AMI + RP booklet: NT	HIV risk-taking behavior	13	0.08 (−0.28, 0.44)	0.05 (−0.31, 0.42)	0.39		
			HIV risk-taking behavior (injecting)	26	0.19 (−0.23, 0.60)	0.08 (−0.34, 0.50)			
Swanson, Pantalon, & Cohen (1999)	Treatment adherence	AMI + ST: ST	Attended FAA (%)	FAA	**0.56ᵃ (0.20, 0.92)**	**0.56ᵃ (0.20, 0.92)**			
Mhurchu, Margetts, & Speller (1998)	Diet & exercise	AMI: ST	Serum cholesterol	13	0.03ᵃ (−0.40, 0.46)	—			
			BMI	13	0.25ᵃ (−0.18, 0.68)	—			
Woollard et al. (1995)	Diet & exercise	AMI + ST: ST	Systolic BP	18	**1.24 (0.80, 1.68)**	**1.04 (0.61, 1.47)**			
			Diastolic BP	18	**1.94 (1.45, 2.44)**	**1.63 (1.16, 2.09)**			
		AMI + PL: PL	Increased physical activity score (%)	12	**0.46ᵃ (0.16, 0.76)**	**0.37ᵃ (0.07, 0.67)**			
				52	0.17ᵃ (−0.12, 0.46)	0.14ᵃ (−0.15, 0.43)			
Harland et al. (1999)	Diet & exercise	AMIᵏ + PL: PL	Increased physical activity score (%)	12	**0.49ᵃ (0.19, 0.79)**	**0.43ᵃ (0.13, 0.73)**			
				52	0.00ᵃ (−0.29, 0.28)	0.02ᵃ (−0.27, 0.31)			
Smith et al. (1997)	Diet & exercise	AMI + ST: ST	Glycemic control	18	0.36 (−0.66, 1.38)	0.20 (−0.82, 1.21)			
Treasure et al. (1999)	Eating disorderʲ	AMI: CBT group	Frequency of binge eating	4	−0.30 (−0.80, 0.19)	−0.30 (−0.80, 0.19)	0.92	53	
			Frequency of vomiting	4	0.31 (−0.18, 0.79)	0.31 (−0.18, 0.79)			

Note. Effect sizes in bold are significant at $p < .05$. TA d is the total attrition effect size. ARP = alcohol-related problems; ASI = Addiction Severity Index; BAC = blood alcohol concentration (peak); BF = brief feedback; BP = blood pressure; CBT = cognitive–behavioral therapy; CC = client-centered counseling; CFB = confrontative feedback; CI = confidence interval; FAA = first aftercare appointment; FTQ = Fagerstrom Tolerance Questionnaire; GGT = gamma glutamyl transpeptidase (a liver enzyme associated with alcohol consumption); NT = no treatment; PL = placebo treatment (e.g., 5-minute interview, informational letter); RAPI = Rutgers Alcohol Problem Index; RP = relapse prevention; RR = risk reduction; SBC = skills-based counseling; SEC = standard ethanol content (per week unless otherwise noted); ST = standard treatment (inpatient or outpatient); TSF = 12-step facilitation. Dashes indicate the data were not available.
ᵃ Calculated from t, F, or chi-square statistic. ᵇ Reported in Bien, Miller, & Tonigan (1993). ᶜ Outpatient sample. ᵈ Aftercare sample. ᵉ AMI including face-to-face feedback. ᶠ AMI including mailed feedback. ᵍ AMI (MI) without feedback. ʰ AMI adapted for brief use by nonspecialists. ʲ Study was not included in the meta-analysis because there was only one study in that problem area. ᵏ One-session AMI.

Table 5
Combined Effect Sizes of Adaptations of Motivational Interviewing (AMIs) by Problem Area

| Problem area | AMIs compared with no-treatment/placebo control | | AMIs compared with active treatment |
	d (95% CI)	TA *d* (95% CI)	*d* (95% CI)
Alcohol (SEC)	**0.25 (0.13, 0.37)**	0.21 (0.09, 0.33)	0.09 (−0.04, 0.23)
Alcohol (BAC)	**0.53 (0.20, 0.86)**	—	—
Smoking cessation	0.11 (−0.05, 0.27)	0.11 (−0.06, 0.27)	—
Drug addiction	**0.56 (0.31, 0.82)**	—	−0.01 (−0.25, 0.25)
HIV-risk behaviors	0.01 (−0.29, 0.31)	−0.01 (−0.30, 0.29)	—
Diet & exercise	**0.53 (0.32, 0.74)**	0.57 (0.33, 0.81)	—
Social impact	**0.47 (0.32, 0.62)**	—	—

Note. Effect sizes in bold are significant at $p < .05$. CI = confidence interval; TA = total attrition; SEC = standard ethanol content (a measure of drinking frequency); BAC = (peak) blood alcohol concentration (a measure of degree of intoxication); dashes indicate the data were not available.

What Is the Sustained Efficacy of AMIs?

The third objective of this meta-analysis was to investigate the sustained efficacy of AMI treatments—that is, the ability of the AMIs to produce lasting symptomatic changes rather than solely an initial response (Westen & Morrison, 2001). Toward this end, the nine studies of AMIs that generated effect sizes at posttreatment and at least one other follow-up point were examined. In these studies, the effect sizes at about 20 weeks posttreatment (0.13, $N = 1,519$) were approximately equal to those at an average of 67 weeks of follow-up ($d = 0.11$, $N = 1,479$), with no significant differences between these two time points, $\chi^2(9, N = 9) = 5.06$, $p > .50$. Note that the mean effect sizes above are not meaningful by themselves, as they confound placebo and active treatment comparison groups (where AMIs have essentially no effect). The basic point is that, regardless of comparison group, the effects of AMIs did not appear to fade significantly over time.

Figure 1. Combined effect sizes of adaptations of motivational interviewing by problem area compared with no-treatment or placebo control groups. SEC = standard ethanol content (a measure of drinking frequency); BAC = (peak) blood alcohol concentration (a measure of degree of intoxication).

What Is the Clinical Impact of AMIs?

Our fourth broad goal in this review was to examine the clinical impact of AMI interventions—the practical value of the effect of AMI to clients or to others with whom clients interact (Kazdin, 1999). We used four different sources of information for this purpose, as displayed in Table 4: (a) percentage improved data, (b) within-treatment effect sizes, (c) pre–posttreatment drinking frequency estimates, and (d) social impact measures.

Percentage improved estimates were available for seven studies of AMIs, six for alcohol problems and one for marijuana dependence. In these studies, 51% of 346 people who received AMI treatments showed noticeable improvement or abstinence (as reported by the original study authors) on substance use measures taken anywhere from 4 weeks to 4 years posttreatment ($M = 53$ weeks posttreatment). Improvement was 54% for stand-alone AMI interventions and 43% for AMIs used as preludes to further clinical services, although this difference was not statistically significant, $\chi^2(1, N = 346) = 3.06$, $p = .08$. Comparatively, 38% of 222 people improved or abstained after no treatment, whereas 35% of 48 people improved or abstained following treatment as usual (e.g., standard inpatient care). Overall, the percentage of people who improved following AMI treatments (51%) was significantly greater than the percentage who improved (37%) with either no treatment or treatment as usual, $\chi^2(1, N = 616) = 10.95$, $p < .01$. When abstinence was the measured target, 38% of 145 clients in AMI groups versus only 18% of 123 clients in no-treatment or treatment-as-usual groups reported abstinence from alcohol or drugs at follow-up, $\chi^2(1, N = 268) = 13.06$, $p < .01$. Thus, administering AMIs in addition to or instead of the usual treatments appeared to improve client success rates from about one third to one half overall, whereas doubling abstinence rates from roughly one in five to two in five.

Within-treatment effect sizes were obtained from 11 studies (across all five problem areas) that used AMIs as stand-alone interventions. The combined within-AMI effect size for these studies was 0.82, a metric that serves as an estimated upper limit of the effects of AMIs before alternative explanations (e.g., passage of time) are ruled out. This means that, on average, a client receiving AMI treatment improved by .82 of a standard deviation on the target outcome measure from intake to posttreatment.
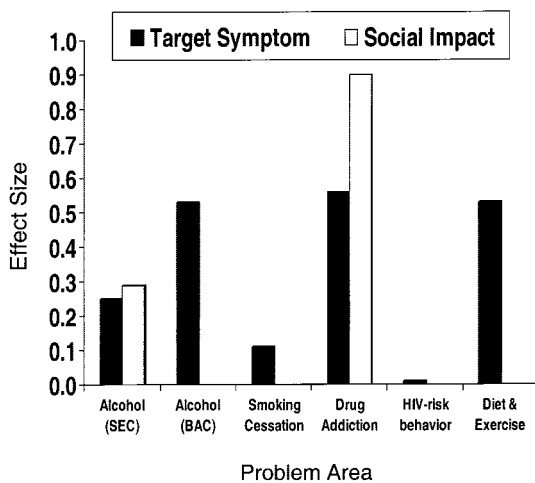
Drinking frequency pre- and post-AMI treatment was estimated from six alcohol studies using stand-alone AMIs. Drinks were measured in SEC units, equivalent to 0.5 oz (15 ml) of absolute alcohol (Miller et al., 1993). On the basis of 373 participants, those receiving AMI treatments lowered their drinking frequency from 35.7 ($SD = 32.9$) SEC drinks/week at pretreatment to 15.6 ($SD = 25.4$) SEC drinks/week at follow-ups of up to 1 year. Whereas this represents a sizeable (56%) decrease in alcohol consumption, the elevated standard deviations must be borne in mind when interpreting these results.

As noted above, combined effect sizes of AMIs were computed separately for social impact measures, variables that tap into important social, occupational, and physical problems related to the target symptom (e.g., absence from work/school, memory loss, legal problems). Displayed in Table 5 (and graphically in Figure 1), these social impact effect sizes were in the medium range overall, reaching 0.29 for alcohol (four studies), 0.90 for drug addiction (two studies), and 0.47 for alcohol and drugs combined (six studies). Interestingly, the social impact effect sizes were not significantly correlated with the target symptom effect sizes from those same six studies, $r(5) = 0.39$, $p > .50$, which suggests that social impact measures may be accessing a different aspect of the overall clinical picture. Thus, AMIs may impact clients in broad and socially relevant ways in addition to bringing about target symptom relief.

What Factors Might Account for Any Observed Differences in Effect Sizes Across These Studies?

Homogeneity Analyses

Using the combined effect size groupings for AMIs compared with control groups and active treatments (when available) across problem areas as a starting point, the tree diagram in Figure 2 illustrates the course taken in subdividing any heterogeneous groupings into smaller clusters until only homogeneous groups (shown with boxes around them) remained. The three studies that compared AMIs with no-treatment or placebo controls for drug addiction initially formed a heterogeneous group, which was subdivided into two homogeneous groups on the basis of whether the study administered a high or a low treatment dose. The study that used a high dose (more than 60 total min of the AMI intervention) produced a larger effect size (0.84, $N = 154$) than the two studies using a low dose of AMI treatment ($d = 0.14$, $N = 96$).

Treatment dose was also used to subdivide the four diet and exercise studies that compared AMIs with no-treatment and/or placebo controls into two groups: one low-dose study ($d = 0.03$, $N = 84$) and three high-dose studies ($d = 0.69$, $N = 282$). This high-dose group was still heterogeneous and therefore was further subdivided into two groups on the basis of MQS, with the low-quality study producing a larger effect size (1.24, $N = 94$) than the two high-quality studies combined ($d = 0.45$, $N = 188$).
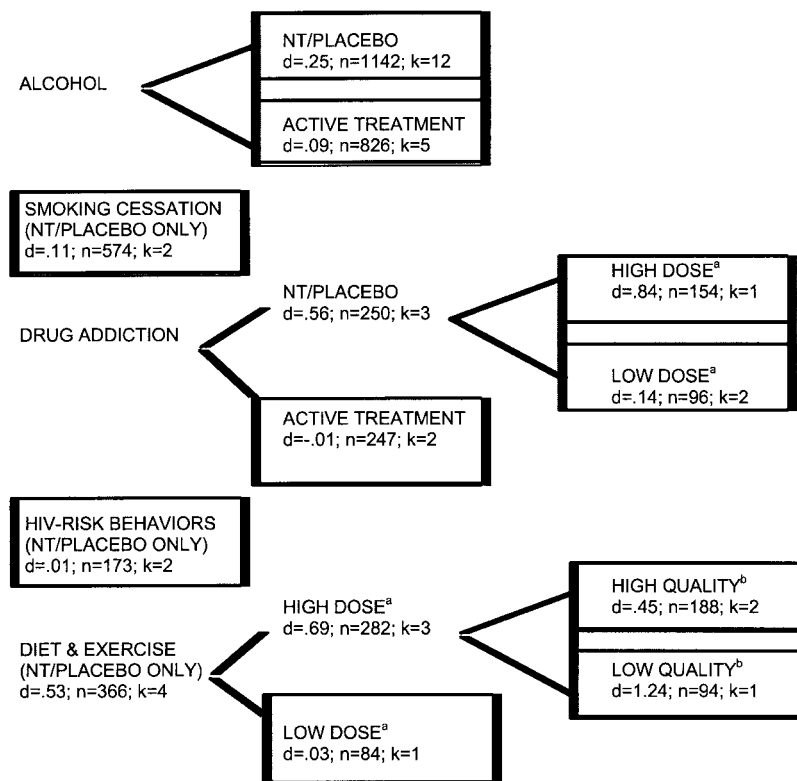

Figure 2. Combined effect sizes of adaptations of motivational interviewing (AMIs) separated by moderator variables into homogeneous groups. Boxes surround groupings of studies that are homogeneous by the Q (chi-square) test. d = combined effect size; k = number of studies; n = number of subjects; NT/PLACEBO = no-treatment or placebo control group. [a] DOSE (total minutes of AMI treatment): low = 5–60; high > 60. [b] Quality (methodological quality score): low < 10; high ≥ 10.

Thus, two categorical moderators, dose and quality, were used to subdivide effect size groupings until all clusters were homogeneous. In two instances, high-dose studies yielded larger effect sizes than low-dose studies, and in one case, two high-quality studies yielded smaller effect sizes than a comparable study of lower quality.

Multiple Regression

Using an exploratory approach, we performed separate multiple regressions for substance abuse studies only and for all studies together. Table 6 presents the final regression model for each analysis. For substance abuse studies, comparative dose ($\beta = .68$, $p < .05$) and format of the AMI ($\beta = .47$, $p < .05$) accounted for 48% of the variance in outcomes. For all studies combined, comparative dose ($\beta = .39$, $p < .05$) and problem area ($\beta = .35$, $p < .05$) accounted for 38% of the variance in outcomes. Other potential moderators (severity of alcohol or drug addiction, methodological quality of the study, and follow-up point) did not significantly predict outcome and therefore were not included in the final regression models.

Accounting for Attrition

Total attrition effect size was estimated for AMI studies compared with no-treatment and/or placebo controls whenever sufficient data were available. The resulting effect sizes (TA d) are displayed for each study in Table 4. Although taking attrition into account generally reduced the effect size estimates, two points are noteworthy here. First, in a few cases, including attrition actually increased the effect size estimates for AMIs (e.g., when effect size was negative). Second, there was only a single case (Aubrey, 1998) in which accounting for attrition actually changed the effect size estimate from significant (nonzero) to nonsignificant for the AMI under study.

As shown in Table 5, we computed combined effect sizes for TA d using the same method as for original d (i.e., weighted by inverse variance). The combined effect sizes were essentially indistinguishable whether or not attrition was taken into account: For alcohol problems, combined TA d (0.21) was slightly lower than combined d (0.25); for diet and exercise problems, combined

Table 6

Summary of Exploratory Regression with Potential Moderators of the Effect Size of Adaptations of Motivational Interviewing (AMIs)

Variable	B	$SE\ B$	β
For alcohol and drug studies ($n = 15$):			
Comparative dose	0.00	0.00	**.68**
Format of AMI[a]	0.14	0.07	**.47**
For all studies ($n = 23$):			
Comparative dose	0.00	0.00	**.39**
Problem area[b]	0.28	0.13	**.35**

Note. Beta weights in bold are significant at $p < .05$. $R^2 = .48$ for alcohol/drug studies and $R^2 = .38$ for all studies.
[a] Format of AMI: 1 = stand-alone treatment, 2 = prelude and/or adjunct to further clinical services. [b] Problem area: 1 = alcohol or drug abuse; 2 = diet and exercise.

TA d (0.57) was slightly higher than combined d (0.53); and for smoking cessation and HIV-risk behaviors, there were no differences between combined TA d and combined d. Combined TA d estimates were not available in the area of drug addiction because of the low number of studies (only one) with individual TA d available.

We also sought to examine data on differential treatment attrition, or whether there were any differences in dropout rates between AMIs and other active treatments. Unfortunately, only one study reported such data (Project MATCH, 1997), and in that study there appeared to be no major differences in attrition between the AMI, 12-step facilitation, and cognitive–behavioral treatments.

File Drawer Problem

We estimated the extent of the file drawer problem and its impact on the conclusions drawn in this review. To reduce the effect size of AMIs in comparison with no-treatment or placebo controls in the areas of alcohol, drug addiction, and diet and exercise problems to a nonsignificant level ($p > .05$), 1,181 filed (or future) studies averaging null results would have to exist. This exceptionally high number bolsters our confidence that the findings presented in the current meta-analysis are resistant to the file drawer threat.

Investigator Allegiance Effects

Studies conducted in Miller's clinic did have a larger mean effect size (0.51) than studies conducted elsewhere (0.21), and this difference was statistically significant, $Q(1, N = 616) = 10.95$, $p < .05$.

Discussion

Summary of Findings

There were 30 clinical trials of AMIs included in this review, representing a wide variety of studies. The prototypical study was conducted in a substance abuse clinic and compared two sessions (99 min) of an AMI with no treatment for about 200 total participants; the study measured target drinking outcomes in SEC per week over about 18 weeks of follow-up, with 74% response rates.

In terms of comparative efficacy, AMIs were equivalent to other active treatments and superior to no-treatment or placebo controls for problems involving alcohol, drugs, and diet and exercise. There was no support for the efficacy of AMIs in the areas of smoking cessation and HIV-risk behaviors in the two studies conducted in each area to date (note: Other studies of HIV-risk behaviors, such as Carey et al., 1997, evaluated group motivational interviewing and hence did not meet inclusion criteria for this review). The efficacy of AMIs for alcohol, drug, and diet and exercise problems was in the medium range overall (ds around 0.50) and appeared to be sustained at follow-up points as long as 4 years posttreatment.

AMIs also showed evidence of clinical impact in these studies. Fifty-one percent of people who received AMI treatment were improved at follow-up compared with 37% of those receiving no treatment or treatment as usual. The average within-AMI effect size was large (0.82), with clients reducing their drinking by 56% from about 36 to 16 standard drinks (SEC) per week as a result of

AMI treatment. AMIs had as much of an effect on social impact measures ($d = 0.47$) as on target symptoms, showing that the treatment could have positive consequences for a wide range of important life problems beyond the target symptoms.

Two categorical moderators, dose and quality, were used to subdivide effect size groupings until all clusters were homogeneous. In two instances, high-dose studies yielded larger effect sizes than low-dose studies, and in one case, two high-quality studies yielded smaller effect sizes than a comparable study of lower quality. Potential moderators were also identified by using exploratory multiple regression. For substance abuse studies, two moderators—comparative dose and format of the AMI—accounted for 48% of the variance in outcomes. For all studies combined, two moderators—comparative dose and problem area—accounted for 38% of the variance in outcomes.

Additional analyses provided evidence that attrition did not pose a major threat to the effect sizes reported in this review. In general, attrition was reasonably low in these studies ($M = 26\%$), especially when it is compared with the average of more than 50% attrition that has been cited in the drug abuse literature (Stanton & Shadish, 1997). As well, taking attrition into account did not alter our combined effect size estimates. Furthermore, it is unlikely that our findings can be attributed to publication bias, as the potential for this bias was mitigated in two ways: (a) we made an effort to retrieve as many unpublished studies as possible in literature searches and by contacting all members of the motivational interviewing network and (b) we estimated that over 1,000 studies with null results would have to be unrecovered to negate the findings of the present meta-analysis, strongly suggesting that these findings are immune to the file drawer threat.

Another factor that may have biased the results of this meta-analysis is the widely studied issue of investigator allegiance (Luborsky et al., 1999). We found that studies conducted in the clinic of the founder of motivational interviewing (W. R. Miller) produced higher effect sizes, on average, than studies conducted elsewhere. There are two possible reasons for this finding: (a) effect sizes may have been inflated in Miller's clinic because of investigator allegiance effects (Luborsky et al., 1999) and/or (b) the superior quality of motivational interviewing training and supervision, as well as ongoing integrity checks, available at Miller's clinic may have resulted in better AMI treatments and hence larger effects. Researchers would do well to emulate the methodology of studies conducted in Miller's substance abuse clinic.

Benchmarks for Comparison

To put this review into perspective, our findings will be compared with other meta-analytic results for alcohol treatments and psychotherapy in general. In the alcohol treatment area, Poikolainen (1999) found that "extended brief interventions" (consisting of about four sessions of treatment, often with a family physician) reduced client drinking by four to five standard drinks per week on average. This estimate, based on eight data sets with over 2,000 total participants, is noticeably smaller than the AMI-assisted reduction of drinking by approximately 20 standard drinks per week reported herein. Additionally, Stanton and Shadish (1997) calculated a combined effect size of 0.38 for family therapy of drug addiction, which is in the same range as the AMI effect sizes

in this review. Unfortunately, no good summary estimates exist at this time for the combined effect size of alcohol treatments, chiefly because of the paucity and poor quality of meta-analyses in this area (Wilson, 2000).

Meta-analytic comparative data are available for psychotherapeutic treatments in general. Lipsey and Wilson (1993) generated the distribution of mean effect sizes from over 300 meta-analyses of psychological, behavioral, or educational interventions, reporting the mean and median effect sizes to be around 0.50 ($SD = 0.29$). The results of our meta-analysis indicate a similar mean effect size of AMIs for alcohol, drug, and diet and exercise problems. On the basis of data from almost 70 meta-analyses, Wampold (2001, p. 70) suggested that the effect size of psychotherapy compared with no-treatment or placebo groups lies within the range of 0.75 to 0.85. Although this value is clearly higher than the AMI effect sizes obtained herein, two points are worth highlighting: (a) some of the AMI treatments included in our review fall more closely into the category of educational or behavioral interventions rather than of psychotherapy per se, so the above comparison to Lipsey and Wilson (1993) may be most appropriate; and (b) the relative treatment doses must be borne in mind. AMI studies featured an average of under 100 min (two sessions) of treatment, whereas most studies included in general psychotherapy meta-analyses delivered at least eight sessions (400+ min) of treatment, four times as much as the AMI studies.

Hence, other psychotherapies have used treatments that are four times as long as AMIs and have yielded combined effect sizes, in previous meta-analyses, that are about 60% larger than the AMI estimates obtained herein. In the current meta-analysis, AMIs were tested directly against other active treatments (e.g., cognitive–behavioral therapy). The mean effect sizes were about zero (0.02), thereby replicating the (in)famous "dodo bird" verdict that all bona fide treatments are roughly equivalent in efficacy (see Wampold et al., 1997). Furthermore, the AMI interventions were shorter than the alternatives in this review by an average of 180 min, yet they still produced similar results. These comparisons of AMIs to longer alternatives (from previous meta-analyses as well as the current one), taken together with our finding of a dose–effect relationship, suggest that longer AMI interventions may generate effect sizes that match those (i.e., around 0.80) generally reported for other psychotherapies.

It is further noteworthy to compare the percentage improved data in this review to analogous data from other meta-analyses. Westen and Morrison (2001) meta-analyzed empirically supported treatments for depression and panic disorder, finding sustained improvement rates of 36%–38% for depression and 50% for panic disorder. Note that both the present meta-analysis as well as Westen and Morrison (2001) used whatever definition of improvement the original researchers reported, which varied substantially from study to study. On the basis of these data, AMI treatments produced client improvement rates for alcohol and/or drug addiction (51%) that were at least equal to rates generated by empirically supported treatments for depression and panic disorder. Moreover, on the basis of over 2,400 clients, Howard et al. (1986) reported that 50% of clients improved in eight sessions of psychotherapeutic treatment—yet AMIs yielded these same improvement rates in an average of just two sessions.

Moderator Findings

A note of caution in interpreting moderator findings in any meta-analysis is warranted in that these are not experimentally derived results; rather, a meta-analysis is a correlational study that emphasizes exploration instead of confirmation (Rosenthal & Di-Matteo, 2001). The search for moderator variables allows one to uncover meaningful patterns in the data that can then be used to formulate potential causal hypotheses to be tested empirically (and experimentally) in subsequent research.

In our overall moderator analysis, problem area was significantly predictive of effect size, such that diet and exercise studies yielded higher effect sizes, on average, than alcohol and drug studies ($\beta = .35$, $p < .05$). This finding is based on only a small number of diet and exercise studies (four) and must remain preliminary. If this proves to be a robust finding in future research, however, there may be a sound reason: Substance use generally involves a physiological addiction, shown through increased tolerance as well as withdrawal symptoms, that may complicate the change process. No equivalent physiological basis exists for diet and exercise problems, which may make these behaviors more amenable to change than entrenched substance use.

Relatedly, this may also be a good place to speculate as to why AMIs did not appear to be efficacious for smoking cessation and HIV-risk behaviors. The most likely possibility at this point is that the findings are not very reliable in these areas: Only four studies total (two in each area) met our inclusion criteria, and both HIV-risk studies came from the same research team. More research is needed before drawing any firm conclusions regarding the efficacy of AMIs for these problems. However, there is one intriguing difference between smoking and some HIV-risk behaviors (e.g., unprotected sex), on the one hand, and alcohol and drug abuse, on the other, that may be worth exploring. The former behaviors are not illegal, whereas the latter behaviors could have serious legal implications (i.e., drinking and driving, underage drinking, using drugs). It is possible that these legal pressures serve to accentuate AMIs in bolstering motivation to change.

We did not find strong evidence for a relationship between methodological quality and study outcome in this review. In the homogeneity analysis, two diet and exercise studies of high methodological quality yielded smaller effect sizes than a comparable low-quality study. However, conclusions based on such small groupings of studies must certainly be regarded as tentative. Furthermore, contrary to our expectations, MQS was not a significant predictor of effect size overall and was therefore not retained in our final regression models. We initially hypothesized that lower quality studies would yield larger effect sizes for AMIs due to systematic errors that could bias the results (e.g., failure to account for treatment dropouts or attrition; no objective or collateral verification of client-reported outcome; nonequivalent comparison groups). Our findings suggest that either this was not the case or, if this effect size inflation did occur, it was counterbalanced by an equivalent deflation as a result of poor AMI treatment delivery (i.e., low fidelity and integrity, discussed in detail later) in these low-quality studies.

The idea that AMIs may be most efficacious as treatment preludes is indeed quite interesting. To our knowledge, it is rare that a treatment can be efficacious both as a stand-alone treatment and as a treatment used adjunctively to enhance the efficacy of a variety of other treatments. Yet this is exactly what some of the research on AMIs suggests, even when the treatments with which they are paired are based on models and techniques quite different from those of the AMIs (Bien, Miller, & Boroughs, 1993; Brown & Miller, 1993). Motivational interviewing was originally developed as a prelude to action-oriented treatment approaches (Miller & Rollnick, 1991), as it is aimed at moving clients toward more advanced stages of change so that the benefit of further clinical services will be amplified. Because of the utility of AMIs as treatment preludes, researchers could consider integrating motivational interviewing components with other therapies for specific clinical problems (e.g., exposure therapy for phobias) to see whether the efficacy of these treatment approaches could also be improved.

In both regression models and in the homogeneity analysis, AMIs showed a significant dose–effect relationship, with higher treatment doses resulting in better study outcomes, as commonly reported in reviews of psychotherapy outcome (e.g., Howard et al., 1986; Shadish et al., 2000). This finding suggests a possible explanation for the increased efficacy of AMIs when used as treatment preludes: Motivational interviewing is designed to prepare clients for change in a small number of sessions, and further sessions may subsequently help clients to initiate and maintain the change. It is feasible that these additional sessions could either consist of other clinical services (e.g., residential treatment) or could consist of simply more AMI treatment. This intriguing possibility merits future empirical investigation in studies using higher doses of AMIs.

Future Directions for Motivational Interviewing Research

As our prior qualitative review indicates (Burke et al., 2002), there has been a real imbalance between internal and external validity in motivational interviewing research, which explains the low quality of some of these studies. It appears that researchers have been more interested in evaluating the extent to which AMIs can be useful with different populations than in constructing rigorous designs to rule out alternative explanations. In most of these clinical trials, training procedures were not carefully described or standardized (e.g., guided by a manual), and integrity checks to measure the implementation of treatment—including videotaping, ongoing supervision, and coding of actual therapist behaviors—were entirely absent from several studies (e.g., Brown & Miller, 1993; Gentilello et al., 1999; Handmaker, Miller, & Manicke, 1999). The first step in enhancing the internal validity of the AMI studies would therefore be to specify more clearly the independent variable of interest—the motivational interviewing treatment under investigation. Future research should therefore include: (a) clearer descriptions of the AMI under study, (b) more careful assessment of treatment fidelity and integrity, and (c) greater uniformity and comparability of AMI treatments across studies.

The results of our meta-analysis suggest another potentially fruitful avenue for AMI research. In this review, AMIs were equivalent to other active treatments but achieved these same effects in considerably less time. Even if AMIs are ultimately no more clinically effective than other interventions, they may well be more cost effective. Although only a handful of AMI studies have examined this important variable to date, cost-effectiveness research would certainly be a valuable addition to the AMI literature

and of special interest to policy makers and clinical administrators alike.

A substantial amount of thought, practice, and research has already been devoted to motivational interviewing, yet we are still far from understanding the precise links between its processes and outcomes (Burke et al., 2002). Although Miller and Rollnick (2002) speculate about possible therapeutic mechanisms, pointing to faith/hope effects, counselor effects (e.g., empathy), and change talk, there is a dearth of evidence about how and why interventions related to motivational interviewing might work. For instance, there is little direct evidence to suggest that AMIs actually work by enhancing motivation or readiness for change. Although AMI clients generally showed an increase in readiness for change following treatment (Handmaker et al., 1999; Mhurchu, Margetts, & Speller, 1998; Treasure et al., 1999), the AMI interventions—with one exception (Butler et al., 1999)—did not appear to *differentially* increase readiness for change in comparison to alternative interventions or controls (Colby et al., 1998; Mhurchu et al., 1998; Saunders et al., 1995; Schneider et al., 2000; Treasure et al., 1999). More process studies are necessary to elucidate the precise mediators and moderators of AMI treatment.

There is another essential question that has eluded any direct answer in the research thus far: What is the efficacy of *motivational interviewing* (as opposed to AMIs)? As noted earlier, there have been no empirical tests of a "pure" form of motivational interviewing, as defined by Miller and Rollnick (1991, 2002), that have not been confounded by feedback or other clinical strategies. At this point, we do not know whether feedback, motivational interviewing, or the combination is essential to produce a therapeutic effect.

In fact, there is some recent empirical evidence suggesting that the feedback component may be more critical than the motivational interviewing component for college student drinkers (Juárez, 2001). Moreover, the results of Schneider et al. (2000) suggest that problem feedback may be efficacious whether delivered in a motivational interviewing or confrontative style. On the other hand, Miller et al. (1993) reported a significant correlation between therapist confrontative behaviors (e.g., challenging, disagreeing, head-on disputes, incredulity, sarcasm) and drinking frequency at 1-year follow-up such that the more the therapist confronted, the more the client drank. Further, Sellman, Sullivan, Dore, Adamson, and MacEwan (2001) found that motivational interviewing was a significantly beneficial addition to problem feedback and outperformed a nondirective counseling feedback condition. To clarify this murky picture, an immediate task for research in this area is to dismantle feedback-based AMIs into their main components—problem feedback and motivational interviewing—so that their relative contributions to outcome can be determined.

Finally, a considerable body of theory and research suggests that motivational interviewing may be effective for clinical areas beyond addictions (Arkowitz & Mann, 2002). Given motivational interviewing's integration of the relationship factors of Rogers (1951) with the active interventions of cognitive–behavioral therapy, this is surely a possibility worth exploring.

References

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

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., rev.). Washington, DC: Author.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Arkowitz, H., & Mann, B. J. (2002). A cognitive–behavioral assimilative integration. In F. Kaslow (Series Ed.) & J. Lebow (Vol. Ed.), *Comprehensive handbook of psychotherapy: Vol. 4. Integrative and eclectic therapies* (pp. 317–340). New York: Wiley.

*Aubrey, L. L. (1998). Motivational interviewing with adolescents presenting for outpatient substance abuse treatment (Doctoral dissertation, University of New Mexico, 1998). *Dissertation Abstracts International, 59*(3-B), 1357.

*Baer, J. S., Kivlahan, D. R., Blume, A. W., McKnight, P., & Marlatt, G. A. (2001). Brief intervention for heavy drinking college students: Four-year follow-up and natural history. *American Journal of Public Health, 91*, 1310–1316.

*Baker, A., Heather, N., Wodak, A., Dixon, J., & Holt, P. (1993). Evaluation of a cognitive–behavioural intervention for HIV prevention among injecting drug users. *AIDS, 7*, 247–256.

*Baker, A., Kochan, N., Dixon, J., Heather, N., & Wodak, A. (1994). Controlled evaluation of a brief intervention for HIV prevention among injecting drug users not in treatment. *AIDS Care, 6*, 559–570.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman & Company.

*Bien, T. H., Miller, W. R., & Boroughs, J. M. (1993). Motivational interviewing with alcohol outpatients. *Behavioural and Cognitive Psychotherapy, 23*, 347–356.

Bien, T. H., Miller, W. R., & Tonigan, J. S. (1993). Brief interventions for alcohol problems: A review. *Addiction, 88*, 315–336.

*Booth, R. E., Kwiatkowski, C., Iguchi, M. Y., Pinto, F., & John, D. (1998). Facilitating treatment entry among out-of-treatment injection drug users. *Public Health Reports, 113*(Suppl. 1), 116–128.

Borkovec, T. D., & Castonguay, L. G. (1998). What is the scientific meaning of empirically supported therapy? *Journal of Consulting and Clinical Psychology, 66*, 136–142.

*Borsari, B., & Carey, K. B. (2000). Effects of a brief motivational intervention with college student drinkers. *Journal of Consulting and Clinical Psychology, 68*, 728–733.

*Brown, J. M., & Miller, W. R. (1993). Impact of motivational interviewing on participation and outcome in residential alcoholism treatment. *Psychology of Addictive Behaviors, 7*, 211–218.

Burke, B., Arkowitz, H., & Dunn, C. (2002). The efficacy of motivational interviewing. In W. R. Miller & S. Rollnick (Eds.), *Motivational interviewing: Preparing people for change* (2nd ed., pp. 217–250). New York: Guilford Press.

*Butler, C. C., Rollnick, S., Cohen, D., Russel, I., Bachmann, M., & Stott, N. (1999). Motivational consulting versus brief advice for smokers in general practice: A randomised trial. *British Journal of General Practice, 49*, 611–616.

Carey, M. P., Maisto, S. A., Kalichman, S. C., Forsyth, A. D., Wright, E. M., & Johnson, B. T. (1997). Using information, motivational enhancement, and skill training to reduce the risk of HIV infection for low-income urban women: A second randomized clinical trial. *Journal of Consulting and Clinical Psychology, 65*, 531–541.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

*Colby, S. M., Monti, P. M., Barnett, N. P., Rohsenow, D. J., Weissman, K., Spirito, A., et al. (1998). Brief motivational interviewing in a hospital setting for adolescent smoking: A preliminary study. *Journal of Consulting and Clinical Psychology, 66*, 574–578.

Dunn, C., DeRoo, L., & Rivara, F. P. (2001). The use of brief interventions adapted from motivational interviewing across behavioral domains: A systematic review. *Addiction, 96*, 1725–1742.

Elkin, E., Gibbons, R. D., Shea, M. T., & Sotsky, S. (1995). Initial severity

and differential treatment outcome in the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *Journal of Consulting and Clinical Psychology, 63,* 841–847.

Garrett, C. J. (1985). Effects of residential treatment on adjudicated delinquents: A meta-analysis. *Journal of Research in Crime and Delinquency, 45,* 287–308.

*Gentilello, L. M., Rivara, F. P., Donovan, D. M., Jurkovich, G. J., Daranciang, E., Dunn, C. W., et al. (1999). Alcohol interventions in a trauma center as a means of reducing the risk of injury recurrence. *Annals of Surgery, 230,* 473–483.

Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research.* Beverly Hills, CA: Sage.

Gleser, L. J., & Olkin, I. (1994). Stochastically dependent effect sizes. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 339–355). New York: Russell Sage Foundation.

*Handmaker, N. S., Miller, W. R., & Manicke, M. (1999). Findings of a pilot study of motivational interviewing with pregnant drinkers. *Journal of Studies on Alcohol, 60,* 285–287.

*Harland, J., White, M., Drinkwater, C., Chinn, D., Farr, L., & Howel, D. (1999). The Newcastle Exercise Project: A randomized controlled trial of methods to promote physical activity in primary care. *British Medical Journal, 319,* 828–831.

*Heather, N., Rollnick, S., Bell, A., & Richmond, R. (1996). Effects of brief counseling among heavy drinkers identified on general hospital wards. *Drug & Alcohol Review, 15,* 29–38.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* San Diego, CA: Academic Press.

Howard, K. I., Kopta, S. M., Krause, M. S., & Orlinsky, D. E. (1986). The dose–effect relationship in psychotherapy [Special issue]. *American Psychologist: Psychotherapy Research, 41,* 159–164.

Hunter, J. E., Schmidt, F. L., & Jackson, G. B. (1982). *Meta-analysis: Cumulating research findings across studies.* Beverly Hills, CA: Sage.

Jessor, R., Donovan, J. E., & Costa, F. M. (1989). *Health Behavior Questionnaire.* Boulder: University of Colorado.

*Juárez, P. (2001). *A randomized trial of motivational interviewing and feedback on heavy drinking college students.* Unpublished master's thesis, University of New Mexico, Albuquerque.

Kaplan, A. (1964). *The conduct of inquiry: Methodology for behavioral science.* Scranton, PA: Chandler.

Kazdin, A. E. (1992). *Research design in clinical psychology* (2nd ed.). Needham Heights, MA: Allyn & Bacon.

Kazdin, A. E. (1999). The meanings and measurement of clinical significance. *Journal of Consulting and Clinical Psychology, 67,* 332–339.

Liddle, H. A., & Dakof, G. A. (1995). Efficacy of family therapy for drug abuse: Promising but not definitive. *Journal of Marital and Family Therapy, 21,* 511–543.

Lipsey, M. W., & Wilson, D. B. (1993). The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis. *American Psychologist, 48,* 1181–1209.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis.* Thousand Oaks, CA: Sage.

Luborsky, L., Diguer, L., Seligman, D. A., Rosenthal, R., Krause, E. D., Johnson, S., et al. (1999). The researcher's own therapy allegiances: A "wild card" in comparisons of treatment efficacy. *Clinical Psychology: Science and Practice, 6,* 95–106.

*Marlatt, G. A., Baer, J. S., Kivlahan, D. R., Dimeff, L. A., Larimer, M. E., Quigley, L. A., et al. (1998). Screening and brief intervention for high-risk college student drinkers: Results from a 2-year follow-up assessment. *Journal of Consulting and Clinical Psychology, 66,* 604–615.

*Martino, S., Carroll, K. M., O'Malley, S. S., & Rounsaville, B. J. (2000). Motivational interviewing with psychiatrically ill substance abusing patients. *American Journal on Addictions, 9,* 88–91.

*Mhurchu, C. N., Margetts, B. M., & Speller, V. (1998). Randomized clinical trial comparing the effectiveness of two dietary interventions for patients with hyperlipidaemia. *Clinical Science, 95,* 479–487.

*Miller, W. R., Benefield, R. G., & Tonigan, J. S. (1993). Enhancing motivation for change in problem drinking: A controlled comparison of two therapist styles. *Journal of Consulting and Clinical Psychology, 61,* 455–461.

Miller, W. R., Brown, J. M., Simpson, T. L., Handmaker, N. S., Bien, T. H., Luckie, L. F., et al. (1995). What works? A methodological analysis of the alcohol treatment outcome literature. In R. K. Hester & W. R. Miller (Eds.), *Handbook of alcoholism treatment approaches: Effective alternatives* (2nd ed., pp. 12–44). Boston: Allyn & Bacon.

Miller, W. R., & Rollnick, S. (1991). *Motivational interviewing: Preparing people to change addictive behavior.* New York: Guilford Press

Miller, W. R., & Rollnick, S. (2002). *Motivational interviewing: Preparing people for change* (2nd ed.). New York: Guilford Press.

*Miller, W. R., Sovereign, R. G., & Krege, B. (1988). Motivational interviewing with problem drinkers: II. The Drinker's Check-Up as a preventive intervention. *Behavioural Psychotherapy, 16,* 251–268.

*Monti, P. M., Colby, S. M., Barnett, N. P., Spirito, A., & Rohsenow, D. J. (1999). Brief intervention for harm reduction with alcohol-positive older adolescents in a hospital emergency department. *Journal of Consulting and Clinical Psychology, 67,* 989–994.

Noonan, W. C., & Moyers, T. B. (1997). Motivational interviewing: A review. *Journal of Substance Misuse, 2,* 8–16.

Poikolainen, K. (1999). Effectiveness of brief interventions to reduce alcohol intake in primary health care populations: A meta-analysis. *Preventive Medicine: An International Journal Devoted to Practice & Theory, 28,* 503–550.

Prochaska, J. O., DiClemente, C. C., & Norcross, J. C. (1992). In search of how people change: Applications to addictive behaviors. *American Psychologist, 47,* 1102–1114.

*Project MATCH Research Group. (1997). Matching alcoholism treatments to client heterogeneity: Project MATCH posttreatment drinking outcomes. *Journal of Studies on Alcohol, 58,* 7–29.

*Project MATCH Research Group. (1998). Matching alcoholism treatments to client heterogeneity: Project MATCH three-year drinking outcomes. *Alcoholism: Clinical and Experimental Research, 23,* 1300–1311.

Rogers, C. R. (1951). *Client-centered therapy.* Boston: Houghton-Mifflin.

Rollnick, S., Heather, N., & Bell, A. (1992). Negotiating behavior change in medical settings: The development of brief motivational interviewing. *Journal of Mental Health, 1,* 25–37.

Rosenthal, R. (1991). *Meta-analytic procedures for social research.* London: Sage.

Rosenthal, R., & DiMatteo, M. R. (2001). Meta-analysis: Recent developments in quantitative methods for literature reviews. *Annual Review of Psychology, 52,* 59–82.

*Saunders, B., Wilkinson, C., & Phillips, M. (1995). The impact of a brief motivational intervention with opiate users attending a methadone programme. *Addiction, 90,* 415–424.

*Schneider, R. J., Casey, J., & Kohn, R. (2000). Motivational versus confrontational interviewing: A comparison of substance abuse assessment practices at employee assistance programs. *The Journal of Behavioral Health Services and Research, 27,* 60–74.

*Sellman, D., Sullivan, P. F., Dore, G. M., Adamson, S. J., & MacEwan, I. (2001). A randomized controlled trial of motivational enhancement therapy (MET) for mild to moderate alcohol dependence. *Journal of Studies on Alcohol, 62,* 389–396.

Shadish, W. R., Navarro, A. M., Matt, G. E., & Phillips, G. (2000). The effects of psychological therapies under clinically representative conditions: A meta-analysis. *Psychological Bulletin, 126,* 512–529.

Shoham, V., & Rohrbaugh, M. (1995). Aptitude × Treatment interaction research: Sharpening the focus, widening the lens. In M. Aveline & D.

Shapiro (Eds.), *Research foundations for psychotherapy research* (pp. 73–95). Sussex, England: Wiley.

Skinner, H. (1982). The drug abuse screening test. *Addictive Behaviors, 7*, 363–371.

*Smith, D. E., Heckemeyer, C. M., Kratt, P. P., & Mason, D. A. (1997). Motivational interviewing to improve adherence to a behavioral weight-control program for older obese women with NIDDM: A pilot study. *Diabetes Care, 20*, 53–54.

Stanton, M. D., & Shadish, W. R. (1997). Outcome, attrition, and family–couples treatment for drug abuse: A meta-analysis and review of the controlled, comparative studies. *Psychological Bulletin, 122*, 170–191.

*Stephens, R. S., Roffman, R. A., & Curtin, L. (2000). Comparison of extended versus brief treatments for marijuana use. *Journal of Consulting and Clinical Psychology, 68*, 898–908.

*Swanson, A. J., Pantalon, M. V., & Cohen, K. R. (1999). Motivational interviewing and treatment adherence among dually-diagnosed patients. *Journal of Nervous and Mental Disease, 187*, 630–635.

*Treasure, J. L., Katzman, M., Schmidt, U., Troop, N., Todd, G., & de Silva, P. (1999). Engagement and outcome in the treatment of bulimia nervosa: First phase of a sequential design comparing motivation enhancement therapy and cognitive–behavioural therapy. *Behaviour Research and Therapy, 37*, 405–418.

Wampold, B. E. (2001). *The great psychotherapy debate: Models, methods, and findings*. Mahwah, NJ: Erlbaum.

Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K., & Ahn, H. (1997). A meta-analysis of outcome studies comparing bona fide psychotherapies: Empirically, "all must have prizes." *Psychological Bulletin, 122*, 203–215.

*Wertz, J. S. (1994). The effect of motivational interviewing on treatment participation, self-efficacy, and alcohol use at follow-up in inpatient alcohol dependent adults (Doctoral dissertation, Virginia Polytechnic Institute, 1994). *Dissertation Abstracts International, 55*(1-B), 6721.

Westen, D., & Morrison, K. (2001). A multidimensional meta-analysis of treatments for depression, panic, and generalized anxiety disorder: An empirical examination of the status of empirically supported therapies. *Journal of Consulting and Clinical Psychology, 69*, 875–899.

White, H. R., & Labouvie, E. W. (1989). Towards the assessment of adolescent problem drinking. *Journal of Studies on Alcohol, 50*, 30–37.

Wilson, D. B. (2000). Meta-analyses in alcohol and other drug abuse research. *Addiction, 95*(Suppl. 3), 419–438.

*Woollard, J., Beilin, L., Lord, T., Puddey, I., MacAdam, D., & Rouse, I. (1995). A controlled trial of nurse counseling on lifestyle change for hypertensives treated in general practice: Preliminary results. *Clinical and Experimental Pharmacology and Physiology, 23*, 466–468.