Published in final edited form as:
J Consult Clin Psychol. 2009 December ; 77(6): 1113–1124. doi:10.1037/a0017189.

From in-session behaviors to drinking outcomes: A causal chain for motivational interviewing

Theresa B. Moyers, **Tim Martin**, **Jon M. Houck**, **Paulette J. Christopher**, and **J. Scott Tonigan**
Center on Alcoholism, Substance Abuse and Addictions, University of New Mexico

Abstract

Client speech in favor of change within motivational interviewing (MI) sessions has been linked to treatment outcomes, but a causal chain has not yet been demonstrated. Using a sequential behavioral coding system for client speech (SCOPE) this study found that, at both the session level and utterance level, specific therapist behaviors predict client change talk. Further, a direct link from change talk to drinking outcomes was observed, and support was found for a mediational role for change talk between therapist behavior and client drinking outcomes. These data provide preliminary support for the proposed causal chain indicating that client speech within treatment sessions can be influenced by therapists, who can employ this influence to improve outcomes. Selective eliciting and reinforcement of change talk is proposed as a specific active ingredient of MI.

Keywords

motivational interviewing; alcohol; mediation; causal chain; client change talk

Motivational interviewing (MI) is an empirically based and widely disseminated psychosocial intervention with applications in the treatment of addictions, health behaviors, and a host of other behavioral concerns. A review of the National Institutes of Health Computer Retrieval of Information on Scientific Projects database and ClinicalTrials.gov website indicates that there are 212 current research projects utilizing or directly investigating the effectiveness of MI, including 114 clinical trials. MI has also found a favorable reception with policy makers who wish to encourage empirically based interventions for public dollars spent on substance abuse treatment. Forty-seven states encourage the use of MI as an example of a preferred treatment for substance abuse and 8 states mandate the use of empirically-based treatment in some fashion, with MI listed as one choice among them. This widespread dissemination has led to a proliferation of MI adaptations as well as a burgeoning industry in training, continuing education, and certification of MI practitioners.

Despite its intuitive appeal to therapists and widespread diffusion, many questions remain concerning this therapeutic method. Some clinical trials have yielded negative results, and the

Correspondence should be sent to: tmoyers@unm.edu.
Theresa B. Moyers, Jon Houck and J. Scott Tonigan, Center on Alcoholism, Substance Abuse and Addictions University of New Mexico.
Paulette Christopher, Department of Psychology University of New Mexico.
Tim Martin now at Kennesaw State University.

Portions of this paper were presented at the Research Society on Alcoholism, June 2007.

behavior of therapists when practicing MI is highly variable both within and between clinical trials despite the use of common materials, training, and supervision (Carroll et al., 2006). What could be causing such variability in both the outcome of the treatment itself and the way it is offered? We argue that this variability is best explained by the presence of one or more unknown active ingredients within MI, which are employed inconsistently by therapists and researchers.

Research to identify and explore active ingredients of MI has only recently begun to attract the attention of the scientific community. One potential active ingredient that has received preliminary research support is that of client speech, in particular the language the client offers in favor of change during an MI session. Since its inception, client speech has been hypothesized as a critical component of MI. Miller (1983) posited that ambivalent clients would decide what they believed about their problematic behavior as they heard themselves discussing it with a therapist, essentially talking themselves into change. From this follows MI's prescription (Miller & Rollnick, 1991; 2002) that therapists should attempt to elicit language in favor of change ("change talk") from their clients during MI sessions, rather than language opposed to it ("counterchange talk"). It is not merely the rote rehearsal of favorable statements that is important, but the sincere and spontaneous emergence of the client's own spoken reasons and desire for making a change within a therapeutic setting. Because their own speech convinces clients what they believe about their behavior, the therapist's active shaping of the client's language during treatment sessions is crucial in directing this active ingredient of the treatment (Miller & Moyers, 2007). Employing this principle, motivational interviewers arrange therapeutic conversations so that the client is offering primarily reasons *for* change whenever possible during treatment sessions. It is this spontaneous, and gently guided, offering of change talk from an ambivalent person, within an empathic therapeutic context, that is one hypothesized mechanism for increasing motivation. Indeed, the recognition and shaping of client speech in this manner is one of the markers of the expert use of MI (Miller & Moyers, 2007).

One specific causal chain for the effectiveness of MI derived from this theory is that the quality of specific therapist behaviors will increase client change talk during treatment sessions, and that this spontaneously emerging change talk will then influence the client's motivation to change behaviors such as drinking, drug use, and management of chronic illness (Moyers et al., 2007). From the perspective of Baron and Kenny's (1986) model for mediation we should first see a relationship between how well therapists employ MI and the amount of change talk that subsequently occurs in treatment sessions (the intervention test). The change talk that occurs during the MI session should then be related to reductions in problematic behaviors (the mediator test). Finally, there should also be a relationship between how well therapists employ MI and subsequent reductions in problematic behaviors (the efficacy test).

Empirical evidence to support the claim that client speech is an active ingredient in MI is sparse but growing, with most attention being paid to the second component of the chain: the relationship between change talk and client outcomes. In general, this research shows a consistent relationship between client language that occurs during an MI session and subsequent desirable behavior change. Researchers have identified two distinct types of change talk that are important in MI: preparatory and commitment language (Miller, Moyers, Amrhein & Rollnick, 2006). Preparatory language is client speech that focuses on a desire, ability, reason or need (i.e, DARN) to change. For example, a client might state "I can never be the kind of mother I want to be while I am smoking crack" (reason) or "If I can quit smoking, I can quit drinking" (ability). Commitment language is a direct statement of intent from the client to change ("I'm going to quit").

In a seminal study, Amrhein and colleagues (Amrhein, Miller, Yahne, Palmer, & Fulcher, 2003) measured the strength of preparatory and commitment language in substance abuse

treatment sessions using a ten point rating scale capturing both positive (committed to change) and negative (committed to the status quo) statements. They found that increased strength of commitment language was associated with more favorable outcomes for drug-using clients, but strength of preparatory language was not. This led Amrhein to emphasize the importance of commitment language an indicator of emerging commitment during MI sessions ands a predictor of MI effectiveness (Amrhein, 2003). Hodgins, Ching, and McEwen (2009) found support for Amrhein's model. In a study of problem gamblers, they found that greater strength of commitment during a motivational interviewing intervention predicted better gambling outcomes twelve months after treatment, though again preparatory elements of change talk were not associated with outcomes. Similarly, Aharonvich, Amrhein, Bisaga, Nunes and Hasin (2008) found that the mean strength of commitment language across cognitive-behavioral treatment for cocaine use was associated with a greater percentage of negative urine screens during treatment.

Other research indicates that both preparatory and commitment language are associated with improved client outcomes. For example, Moyers et al. (2007) examined treatments sessions from all three treatment conditions in Project MATCH (Project MATCH Research Group, 1997): Twelve Step Facilitation Therapy, Cognitive Behavioral Coping Skills Therapy and Motivational Enhancement Therapy. Using an early version of the Motivational Interviewing Skills Code (MISC 1.0: Miller, 2000), the frequency of client change talk was measured without differentiating between preparatory and commitment categories. Using this simpler measure, the authors found that higher frequencies of change talk predicted fewer drinks per drinking day at the distal follow up point, even in these three very different treatments. Baer and colleagues (2008) evaluated adolescent language during a brief motivational intervention with substance abusing homeless youth. They found that client change talk focused on reasons for change was associated with significantly lower rates of substance use at the one month follow up point. Furthermore, client statements refuting a desire or ability to change predicted greater substance use at both the one and three month outcomes. Commitment language was not associated with outcomes. Other researchers have found that client speech focusing on ability (Gaume, Gmel, & Daeppen, 2008), and action orientation (Strang & McCambridge, 2004) have predicted drinking and marijuana use after MI interventions.

Although there is some disagreement about the type of client speech that is most important (change talk vs. commitment language) and the way it is measured (frequency vs. strength), these studies uniformly support client change talk as a predictor of benefit from MI sessions. These studies support the hypothesis that change talk is an active ingredient in this treatment; however, the nature of the data do not permit us to rule out alternative explanations for the relationship between change talk and outcomes. It is possible, for example, that highly motivated clients exhibit more change talk within sessions, whereas less motivated clients simply exhibit less of it. Change talk, then, may be the bellwether for eventual change without having any causal relationship to it in much the same way that smoke indicates fire but does not cause it.

Assuming that change talk does predict treatment outcomes, do therapists have any influence over it in treatment sessions? This question is critical, since an active ingredient of a psychological treatment should be amenable to the influence of the therapist. Catley and colleagues (2006) used the MISC 1.0 to evaluate therapist behaviors and client speech in a clinical trial employing MI to reduce smoking. They found that MI Consistent behaviors on the part of the therapist were associated with higher levels of change talk. Further, higher levels of MI Spirit (a global rating) also predicted higher levels of client change talk. Houck and Moyers (2008) also used the MISC 1.0 to evaluate the relationship between therapist behaviors and client speech in audiotaped therapy samples from a training study in which the MI skills of 140 substance abuse therapists were evaluated before and after workshop training (Project

EMMEE: Miller, Yahne, Moyers, Martinez, & Pirritano, 2004). They found that even before MI training, there was a significant relationship between MI consistent therapist behaviors and client change talk. The magnitude of this relationship became stronger after training, when MI skills increased, and persisted through the 12-month follow-up period.

Moyers and Martin (2006) approached this question by evaluating the temporal relationship between therapist behaviors and client speech, reasoning that if MI Consistent behaviors do elicit change talk there should be a sequential relationship between the two in therapy sessions. That is, occasions of appropriate MI behavior (such as emphasizing the client's autonomy) should be more likely to be immediately followed by change talk than are occasions of inappropriate MI behavior (such as confrontation). The authors adapted MISC coding system so that the temporal order of behaviors could be preserved. This new coding system, the Sequential Code for Observing Process Exchanges (SCOPE: Martin, Moyers, Houck, Christopher, & Miller, 2005) was used to evaluate audio recordings of 38 therapy sessions from the Project MATCH Motivational Enhancement Therapy (MET) condition. The authors found that the likelihood of client change talk was higher following behaviors consistent with MI (MICO) and lower following therapist behaviors inconsistent with MI (MIIN). In addition, MIIN behaviors on the part of the therapist were more likely to be followed by client counterchange talk. These transition probabilities, because they preserve the temporal relationship between the process variables, provide stronger support for a causal hypothesis than would be found in a correlational design. Gaume, Gmel, Faouzi, and Daeppen (2008) also used a sequential coding system to evaluate transition probabilities between MI Consistent therapist behaviors and client change talk. Replicating the Moyers and Martin (2006) study, they found that MI Consistent therapist behaviors were associated with increased client change talk during an emergency room intervention for problem drinkers.

Despite such encouraging research to support change talk as a potential mediator in motivational interviewing, there has not been a study which examines all elements of a mediational model within a single intervention. The current study focuses on exploring the hypothesized causal chain between therapist behaviors, client speech, and substance abuse outcomes using a large, representative sample from the Project MATCH archive of MET tapes and employing a sequential coding system (SCOPE) that allows the generation of transition probabilities among these variables. We hypothesized that if MI is rendered effective by therapists' elicitation of client change talk, then the probability of change talk in these MET sessions would be greater when therapists used behaviors prescribed in MI. Furthermore, we hypothesized that the change talk that occurred in these MET sessions should be associated with improved substance abuse outcomes.

Methods

Data Source

All procedures for this study were approved and overseen by the Institutional Review Board of the University of New Mexico. We obtained all available Motivational Enhancement Therapy (MET; n = 225) tapes in the Project MATCH archives from five sites where IRB committees were willing to give permission for secondary analysis of the therapy sessions. The tapes we obtained included Motivational Enhancement Therapy sessions from both the outpatient treatment and the aftercare treatment arms of Project MATCH. Of these tapes, 118 were audible Session 1 tapes and comprised the data pool for this study. We elected to use only Session 1 tapes because they comprised the largest population of independent observations. Of the remaining 107 tapes, 32 contained significant others and could not be coded with the SCOPE. Another 12 tapes were inaudible, incorrectly labeled, or incomplete. The remaining 63 tapes were of subsequent MET sessions.

Drinking Outcome Variables

We obtained drinking outcome data for these participants from the Project MATCH data set, matching outcome to therapy session via the client identification number. Outcome variables derived from this data set were weekly measures of Percent Days Abstinent (PDA) and Drinks per Drinking Day (DDD) for baseline and in-treatment follow up points. We also created a composite measure for each week, Drinks per Week (DW), by multiplying DDD × [7 × (1-PDA)]. These outcome measures are obtained from the Form 90 (Miller & Del Boca, 1994), a series of interviews that combines both a timeline follow-back method and a grid averaging method of reporting drinking behavior. The Form 90 has been shown to have relatively high test-retest reliability (Tonigan, Miller, & Brown, 1997). Within Project MATCH, self-reported drinking as measured by the Form 90 was consistent with physiological measures of drinking, and generally higher than that obtained from collateral interviews (Project MATCH Research Group, 1997). Descriptive statistics on these outcome measures and all covariates can be found in Table 1. MET sessions in Project MATCH occurred at weeks one, two, six, and twelve. Inspection of the data revealed strong nonlinearity in the pattern of DW over time, specifically that drinking declined after each of the four MET sessions, then rose somewhat until the next session. Because of this we elected to examine the period from baseline through week five, when the effects of the initial MET session would be most clear. This includes the first and second MET sessions, both of which occurred within the first two weeks of treatment. Both linear and quadratic slopes were evaluated for this period, with the intercept (initial status) of both slopes defined as average DW in the fifth week of treatment (i.e., baseline = −4, week one = −3, week five = 0). Abstinent participants were included in this model. Because DW was not normally distributed, a Poisson model with log link was used in all multilevel models. A zero-inflated Poisson model was not used in multilevel modeling as it is not available under the generalized linear model framework. However, the results for individual follow-up points modeled using a zero-inflated count regression (SAS proc countreg) generally mirrored those seen in the multilevel model.

Baseline Measures of Client Self-Efficacy, Alcohol Involvement, and Readiness to Change

Self-efficacy, alcohol involvement, and readiness to change are all predictors of substance use in general and specifically in this sample, so all were considered for inclusion as covariates in this model. (e.g., Project MATCH Research Group, 1997). The Alcohol Abstinence Self-Efficacy scale (AASE: DiClemente, Carbonari, Montgomery, & Hughes, 1994) is a 20-item measure that was designed to capture the social psychological construct of self-efficacy as described by Bandura (1989), with respect to drinking behavior. Items are rated on a scale of 1 (“not at all confident” to 5 (“very confident”). It captures both temptation to drink and self-efficacy to abstain from drinking. Internal consistency ranges from .92 to .95, with a negative correlation between temptation and self-efficacy ranging from −.58 to −.65 (DiClemente et al., 1994; DiClemente & Hughes, 1990). In this study, we use the total AASE-C score assessed at intake with a possible range of 20–100. Higher scores indicate higher self-efficacy.

The University of Rhode Island Change Assessment scale (URICA; McConnaughy, Prochaska, & Velicer, 1983) was developed to measure attitudes related to each of the stages of change. The Readiness subscale of the URICA was calculated by adding the Contemplation, Action, and Maintenance subscale scores and subtracting the Precontemplation subscale score (Carbonari, DiClemente, & Zweben, 1994; Connors, Tonigan, & Miller, 2001). The subscale has a possible range of 0–14, with higher scores indicating greater readiness to change.

The Alcohol Involvement measure used was ALCINVOL, a third-order scale from the Alcohol Use Inventory (AUI; Wanberg, Horn, & Foster, 1975). This scale is a broad indicator of the involvement of alcohol in an individual’s life, and is based on factor analysis of the six second-order scales of the AUI. It is the sum of 42 items from the 224-item parent measure. The scale

ranges from 1 to 68, with higher scores indicating greater alcohol involvement. Internal consistency of this scale in the MATCH dataset was .93 (Rychtarik, Miller, & Tonigan, 2001). Criterion and construct validation of ALCINVOL are also acceptable for research purposes (Skinner & Allen, 1983).

MET Therapists

To qualify for inclusion in Project MATCH, therapists had to meet five eligibility criteria: completion of a master's degree or higher in psychology or a closely related field; at least two years of clinical experience; submission of a work sample for review; experience with the treatment condition; and experience treating substance abuse. All therapists included in Project MATCH received training in the background and rationale for the project, as well as specific training for their assigned treatment condition. Therapists in Project MATCH were nested within treatment; i.e., MET therapists performed only MET sessions, not Twelve Step Facilitation Therapy (TSF) or Cognitive Behavioral Coping Skills Therapy (CBT). A total of 26 therapists were trained and certified for the MET condition of Project MATCH; the sample of available tapes included sessions from only 13 of these therapists. This subset of 13 therapists represents all MET therapists from the MATCH sites where Institutional Review Boards allowed secondary analysis of session tapes for this study. Detailed information on therapist training and supervision in Project MATCH can be found elsewhere (Carroll et al., 1998; Project MATCH Research Group, 1998; Witte & Wilber, 1997).

Behavioral Coding System

To evaluate these therapy sessions, we used the SCOPE (Martin et al., 2005). The SCOPE is a mutually exclusive and exhaustive systematic observation system; that is, every audible utterance is coded, and is assigned one and only one code. The system is designed to evaluate the verbal behavior of both therapist and client during MI therapy sessions (Moyers & Martin, 2006). The SCOPE is used to generate transition probabilities between therapist and client behaviors within MI therapy sessions. These transition probabilities describe the probability that a behavior, such as client change talk, will occur immediately after another behavior, such as a therapist reflection, has occurred. In addition to these transition probabilities, the SCOPE also produces frequency counts for specific therapist and client behaviors, analogous to those produced by the MISC 2.0.

Two separate passes are required for SCOPE coding, and transcripts are always used in conjunction with audio recordings. In the first pass, a coder breaks down ("parses") the transcribed therapy session into sequentially numbered utterances. Within the SCOPE coding system, an utterance is a segment of speech that represents a complete thought. In the second pass, a different coder assigns a behavioral code for each utterance. Six individuals provided SCOPE coding for this project, with three serving through the life of the project. Three coders were graduate students at the University of New Mexico, two were undergraduates and one coder was a staff member with a BA in Psychology. None of the coders were therapists or enrolled in clinical training programs. Training of coders to acceptable levels of reliability required 60 hours over six weeks time. Information regarding the use of the SCOPE for coding MI sessions, as well as protocols for training coders and establishing reliability can be found elsewhere (Moyers & Martin, 2006). The SCOPE manual and training materials are available for download at http://casaa.unm.edu/codinginst.html.

Reliability

Previous experience with sequential coding (Moyers & Martin, 2006) convinced us that having coders both parse and code a transcript simultaneously led to insurmountable difficulties in assessing reliability. Specifically, when there were disagreements in parsing, we could not easily align sequences of utterances to assess category agreement for coders. To address this,

we modified our procedure so that all transcripts in this study were parsed in a separate review of the tape. Different coders were used for the parsing and coding passes to avoid a bias in parsing.

Four transcripts were randomly selected for double-parsing, yielding an average number of 7,205 words. We followed Bakeman and Gottman's (1997, pp. 68–69) suggestion to treat each word as a possible breakpoint between utterances. For each double-coded parsing transcript the word count at which each utterance was demarcated was noted for both coders. Cohen's kappas (Cohen, 1960) were computed for each transcript pair, yielding an estimate of parsing reliability. Three of the six coders parsed transcripts.

Reliability of utterance- to-utterance agreement among coders—A reliability sample of 40 randomly-selected tapes was coded by all three primary coders. In addition, the three remaining raters double-coded a subset of the reliability sample as they were available to do so. Personnel changes precluded these coders from completing all tapes in the reliability sample.

Categorical agreement among coders was estimated using Cohen's kappa. Bakeman and colleagues (Bakeman, Quera, McArthur, & Robinson, 1997) have shown that kappa is affected by the number of categories and the distribution of base rates among categories. Specifically, at a given level of coder accuracy, kappa is inflated when categories are many but suppressed when categories have widely different base rates. Bakeman et al. (1997, fig. 1) suggests that for a relatively small coding system of 10 categories and widely differing base rates among those categories, coders who are 80% accurate will attain a kappa of approximately .55.

Reliability of frequencies for behavior categories—Reliability for frequencies was estimated with the intraclass correlation coefficient (ICC; Shrout & Fleiss, 1979). This statistic can be computed for two or more variables. Thus, pairwise reliability was estimated for each pair of coders as well as groupwise reliability across all coders. We used a mixed model ICC, with items as fixed factors and raters as random factors. According to Cicchetti (1994), intraclass correlations below .40 are poor, those from .40 – .59 are fair, from .60 – .74 are good, and those from .75 – 1.00 are excellent.

Reliability of transition probabilities—Estimating reliability for transition probabilities is less straightforward. Traditionally, inter-rater agreement for sequential coding schemes has been estimated by aligning the sequences of two coders, cross-tabulating them to generate an agreement matrix, and computing percent agreement and Cohen's kappa (Cohen, 1960). Bakeman et al. (1997) pointed out that although using kappa in this manner allows measurement of categorical agreement, it does not adequately capture agreement about the sequential pattern of the codes (transition probabilities), which is typically of central interest when using sequential observation systems. Instead, they recommended computing an (ICC) for Yule's *Q*, to determine the statistical significance of transition probabilities. Yule's *Q* statistic is a modified odds ratio for a given code, which ranges from −1 to +1, making it analogous to a linear correlation coefficient. In practice, for a given code A, all codes are coded as A or Not-A for each coder, and these two series are cross-tabulated. If f_1 is the frequency of positive matches (both coders agree that an utterance is A), f_2 is the frequency of disagreement where coder 1 coded the event as A, f_3 is the frequency of disagreement where coder 2 coded the event as A, and f_4 is the frequency with which both coders agreed that the event was Not-A, Yule's *Q* statistic is then defined as

$$Q=\frac{f_1 f_4 - f_2 f_3}{f_1 f_4 + f_2 f_3}$$

While Yule's *Q* will capture the significance of transition probabilities, it cannot be used when the frequency of a particular transition is 0. This is often the case when a coding system has many categories or infrequently occurring categories, as with the SCOPE. The most important conclusions that are drawn from transition probabilities are based on their direction (more or less probable than predicted by chance) and whether they are statistically significant. Therefore, in addition to the ICC for Yule's *Q*, we assessed the reliability of transition probabilities by dummy coding each transition as significant or non-significant (1 or 0, respectively) for each coder, and computing the kappa of these cross-tabulated codes.

Analysis Plan

Predicting client speech from therapist behaviors with sequential analysis—To characterize the sequential patterns of therapist-client interactions, we computed transition probability matrices from the SCOPE codes using the GSEQ application (Bakeman & Quera, 1995). Intuitively, a transition probability is the probability that a behavior B will occur, given that a behavior A occurred previously. More formally, a transition probability is defined as

$$P(B_{i+j}|A_i)=\frac{P(A_i \cap B_{i+j})}{P(A_i)}$$

where B is event B, A is event A, and i and j denote sequential position.

A total of 60,856 transitions were observed, or an average of 516 per session. Some very low frequency categories were collapsed in order to obtain reasonable expected cell frequencies (Wickens, 1982). Categories were combined according to both theoretical and practical concerns (Miller, 2000). For the therapist, we were most interested in behaviors specific to MI practice. Consequently, we combined Affirm, Emphasize Control, Permission Seeking and Support into the category Sequential MI-consistent (sMICO). As in our previous work (Moyers & Martin, 2006), this measure differs from that described in the MISC manual (Miller, 2000) in that sMICO does not include therapist reflections or open questions. We separated these codes in order to examine separately the effects of questions and reflections on subsequent behaviors. Behaviors that are contraindicated in MI, including Advise, Confront, Direct, and Warn, were combined into MI-Inconsistent category (MIIN). The measure of MIIN used throughout this paper is identical to the original description (Miller, 2000). Therapist questions and reflections were each combined and analyzed in distinct ways that emphasize aspects of the coding system. The valence of questions was analyzed by collapsing across the open vs. closed dimension to create the categories of Question Positive aspects of the target behavior (QPOS), Question Negative aspects of the target behavior (QNEG), and Question Neutral with respect to the target behavior (QNEUT). Likewise with reflections we collapsed across the simple vs. complex dimension to emphasize valence with Reflect Change Talk (RCT), Reflect Counterchange Talk (RCCT), and a single category that encompassed Reflect Both and Reflect Neutral, which we refer to as simply REF. The reason that we combined the Reflect Both and Reflect Neutral categories was that in practice the Reflect Both category was so rare that it was statistically unusable.

In order to increase statistical power, data were pooled across participants to obtain population parameter estimates of the transition probabilities. Because this method could be influenced

by extreme cases (i.e., particularly long sessions or otherwise deviant interactions), we also estimated the transition matrices for each session, and assessed the frequency of statistically significant transitions for certain transitions of interest. These included the transitions from the therapist categories of Question, Reflection, sMICO, and MIIN to the client categories of CT and CCT.

Predicting client speech from therapist behaviors—In order to assess the magnitude of the relationship between therapist behaviors and client speech, we conducted a series of parallel regressions using MICO and MIIN as predictors of client speech. Specifically, client speech was collapsed into two global categories consistent with previous research protocols: Change Talk (CT), and Counterchange Talk (CCT). CT combined all of the statements of commitment, desire, ability, reason, need, taking steps, and "other" that favored change, while CCT is simply the combination of all of these statements that favored the status quo. Similarly, we used therapist behaviors to predict percent change talk (PCT; i.e., CT/(CT+CCT)), a summary measure described in MISC 1.0 (Miller, 2000). Within this and all subsequent analyses we used the original definition of MICO, which includes open questions and all types of reflection.

Predicting client outcome from client speech—Multilevel modeling (Hox, 2000; Kreft & de Leeuw, 1998; Raudenbush & Bryk, 2002) was used to assess the ability of client speech to predict drinking outcome (Drinks per Week; DW) while accounting for the hierarchical structure of the data. All multilevel modeling was done using HLM version 6 (Raudenbush, Bryk, & Congdon, 2004). A two-level model was used, with repeated measures of DW in the first level nested within participants in the second level. Units in the first level were six weekly measures of DW from baseline through the end of week five of treatment. Change talk was entered in the second level, with alcohol involvement, alcohol abstinence self-efficacy, and readiness for change entered as covariates. Although therapist effects were expected since clients had been nested within therapist in Project MATCH (Project MATCH Research Group, 1998), inclusion of a third level to account for therapist effects did not explain a significant proportion of variance (0.2%, $p > .50$). Thus the two-level model was retained.

Predicting client outcome from MI consistent therapist speech—Multilevel modeling was also used to assess the ability of therapist behaviors to predict drinking outcome (Drinks per Week; DW). As with the client speech model, a two-level model was used, with repeated measures of DW in the first level nested within participants in the second level. In this model, behaviors consistent with motivational interviewing (MICO) were entered in the second level, with alcohol involvement, alcohol abstinence self-efficacy, and readiness for change entered as covariates.

Mediation test—A diagram of the proposed mediation model is given in Figure 1. In this study, the independent variable (X) and intervening variable (M) occur at the same level, therefore this multilevel mediation model is conceptually similar to the model described by Baron and Kenny (Krull & MacKinnon, 2001;MacKinnon, 2008). We used the procedure described by MacKinnon and colleagues (MacKinnon, Fritz, Williams, & Lockwood, 2007) to produce confidence intervals for the mediated effect using the supplied SAS macro. This procedure is analogous to the product of coefficients method described by Sobel (1982), but does not assume that the product of coefficients is normally distributed (Preacher & Hayes, 2008;Williams & MacKinnon, 2008). Instead, confidence limits are computed based upon the shape of the observed distribution. Simulation studies have demonstrated this method to have considerably more power than not only the Baron and Kenny steps and Sobel test (MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002) but also the percentile bootstrap and bias-corrected bootstrap methods (MacKinnon et al., 2007).

Results

Reliability Estimates

Reliability of parsing—Coders achieved an average kappa of .89, with a range of .87 to .91, indicating good agreement as to what constituted an utterance.

Reliability of utterance- to-utterance agreement among coders—The mean Cohen's kappa was .75, with a range of .56 – .87. All reliabilities for this data set were at or above the minimal acceptable values suggested by Bakeman et al. (1997) for a system with many categories and wide differences in the base rates of categories.

Reliability of frequencies for behavior categories—Intraclass correlations for the frequencies across coders all individual behavior counts are available as online supplementary material. Intraclass correlations for the frequencies across coders within summary scores are presented in Table 2. Reliability estimates for our summary scores range between .49 and .98. In general, it appears that coders in this study had more difficulty coding questions about the negative aspects of current drinking. They also had difficulty coding complex reflections as well as those reflections in which both positive and negative content about drinking were included. Of our 16 summary scores, 13 are in the good to excellent range.

Reliability of transition probabilities—The ICC for Yule's *Q* for the three coders who coded the entire reliability sample was .54, indicating that our coders characterized the temporal pattern of this data with fair reliability. Cohen's kappas for the significance of the transition probabilities ranged from .70 to .72. Given the work of Gardner (1995) and Bakeman et al. (1997), we would expect lower estimates of agreement for the transition probabilities than for the individual categories, and that is the pattern that we obtained. We examined the transitions that were disputed among the coders and found that there were no cases in which one coder found the transition to be more probable than chance while another coder found that same transition to be less probable than chance.

Regression analyses

Predicting client speech from therapist behavior—Table 3 presents the regressions of CT and CCT on MICO and MIIN. MICO was a significant predictor of CT ($\beta = 0.47$, $p < .001$), such that higher frequencies of MICO were associated with higher frequencies of CT. Both MICO ($\beta = .27$, $p < .01$) and MIIN ($\beta = .23$, $p < .01$) were significant predictors of CCT indicating that higher frequencies of either were associated with higher frequencies of CCT. The overall model for percent change talk (i.e., CT/(CT+CCT)) only approached significance ($p = .054$).

Multilevel Analyses

Predicting client drinking outcomes from client speech—The results of this analysis are presented in Table 4. Neither alcohol involvement (AIM) nor self-efficacy (AASE) were significant in any model and were therefore dropped as covariates.

There were significant effects in both the first and second levels. Change talk predicted drinking status at week five, with more change talk associated with fewer drinks per week. In addition, change talk predicted both the linear and quadratic slopes of DW during the treatment period. More change talk was associated with fewer drinks per week at week five, and also with the linear and quadratic slopes of drinking from baseline through week five.

Predicting client drinking outcomes from therapist behaviors—The results of this analysis are presented in Table 5. As with the analysis of client speech, linear and quadratic slopes were examined in a Poisson model with log link.

There were significant effects in both the first and second levels. The frequency of MICO behaviors predicted drinking status at week 5, with more MICO behaviors associated with fewer drinks per week. In addition, MICO predicted the quadratic slope of DW from weeks 0 through five of the treatment period. Higher levels of MICO were associated with fewer drinks per week at week five, and also with quadratic slope of drinking from baseline through week five.

Mediation analysis—The results of mediation path c′ are presented in Table 6. Although the MICO coefficient changed when change talk was added to the model, this difference was not significant when evaluated using a chi square difference test ($\chi^2_{diff}(3)=1.432$, $p > 0.5$). However, the confidence limits produced by testing the *a* and *b* paths using the more sensitive PRODCLIN procedure (MacKinnon et al., 2007) indicated a significant indirect effect (99% CI: .001, .002) using the asymmetric distribution of the product method. Within this sample, change talk appears to mediate between MICO and drinking outcomes, accounting for about 30% of the effect. A second test evaluated both change talk and counterchange talk as mediators between MICO and drinking outcomes using a SAS macro designed to test three-path mediation models (Taylor, MacKinnon, & Tein, 2008), modified to model a Poisson-distributed outcome. This model was not significant under any of three different testing methods (Bias-corrected bootstrap 95% CI: −.012, .001).

Sequential analyses

Predicting client speech from therapist behaviors—Figure 2 illustrates the relationship between CT and two therapist behaviors of theoretical interest, RCT and questioning negative aspects of the target behavior (QNEG). For example, it can be seen that when therapists offered QNEG, the probability of CT in the next utterance was .49, and when clients offer CT, the probability of RCT in the next utterance was .17. Complete transition probabilities are presented in Table 7. In the table, behaviors that occurred first (the given behaviors) are in rows, while subsequent behaviors are in columns. For example, to find the probability that the client offered CT, given that the therapist uttered a reflection of change talk (RCT), one would begin by finding the row corresponding to RCT, and then the column corresponding to CT (in this example, the probability is 0.44). From the table, it can be seen that CT was more likely than expected by chance following CT, CCT, QPOS, QNEG, and RCT. Clients were less likely than expected by chance to express change talk following FN/ASK, MIIN, RCCT, REF, or other therapist behaviors. CCT was likely to follow CT, CCT, QPOS, QNEG, or RCCT. CCT was unlikely to follow FN/ASK, MICO, or other therapist behaviors.

Discussion

The findings in this study support our a priori hypothesis that client change talk, once mobilized by therapist behaviors during MI sessions, would lead to reduced drinking. Furthermore, this relationship was evident above and beyond the impact of client readiness to change, which was statistically controlled for in the model. Finally, it appears that the more general category of change talk, which includes both preparatory and commitment language, can predict the success of MI sessions, rather than only commitment language as reported by Amrhein et al. (2003) and Hodgins, Ching and McEwen (2009). Our data imply that, at least for drinking behavior, therapists should work to elicit and reinforce all types of client statements in support

of change rather than according preferential status to only those indicating commitment or intent.

The implications of our data for the provision of clinical services are specific and concrete: to obtain better outcomes using MI, clinicians should attend carefully to client language about change and use strategies recommended in the practice of MI (asking questions likely to result in change talk, reflecting change talk when it occurs and emphasizing client choice) to gain more of it. A strength of this study is that it also indicates what clinicians ought *not* to do if they wish to hear change talk. That is, they should avoid confrontation, giving advice or raising concerns without permission, and telling clients directly what to do. The schism between these data and the most commonly used practices in current substance abuse treatment could not be more obvious.

Our study revealed some features of change talk that were surprising. One surprise was the striking amount of ambivalence (back-and-forth-ness) in client speech, even when change talk was offered. Far from being a straightforward progression of desire, ability, reason, need and commitment to change that culminated in a plan, these Project MATCH clients, even those who would soon make a change in their drinking, offered change talk embedded within a web of what would usually be called denial and minimization. Change talk, although predictive of better outcomes, occurred nearly simultaneously with counterchange talk. This was so common that our coders learned to listen for a "change talk sandwich" – an utterance of change talk surrounded by two instances of counterchange talk ("I don't need to change my drinking. Well, I mean, I need to cut down for sure, but no way am I problem drinker like they say I am."). Furthermore, both MICO and MIIN were associated with an increased probability of counterchange talk indicating that counterchange talk may persist, even when the therapist is appropriately affirming, supportive and careful to emphasize the client's autonomy, and even while change talk is simultaneously increasing in frequency. We conclude that important opportunities for eliciting motivation often lie embedded within client resistance in MI sessions. Wise therapists will not suppress or challenge client statements that appear to resist change, but rather view them as a normal context for eliciting and reinforcing the more favorable speech that is likely to be surrounding them.

Another surprise in our data was the relative weakness of the hypothesized relationship between MI Consistent behaviors and subsequent client change talk, especially compared to the much stronger link between reflections and change talk. These transition probabilities indicate that reflective listening skill may be more effective in eliciting change talk than are supporting autonomy, affirming, or deferring advice. Further, it appears that therapists who wish to see more change talk should selectively reflect the change talk they hear, and provide fewer reflections for counterchange talk. What therapists reflect, they will hear more of. These data suggest that the relatively sophisticated MI skill (Miller & Moyers, 2007) of differential reinforcement of client speech may be more important in evoking high levels of change talk than previously supposed. However, because the MICO variable in our multilevel model was a composite of several therapist behaviors, it is unclear whether a single therapist behavior, or some additive or multiplicative combination, was driving this effect.

While our results are generally consistent with those of Amrhein et al. (2003), there are several points of departure. Previous efforts in our research group to attain reliability on strength ratings for client language have repeatedly failed. Therefore, we made the pragmatic decision to focus upon reliable ratings of change talk categories (DARN-C). Using frequency of CT allowed us to calculate utterance by utterance reliability among our coders; we were able to determine exactly which utterance coders were rating and what code they are assigned to it for each and every client utterance. Amrhein's system relies on mean strength ratings for entire deciles, meaning that coders are not necessarily rating the same utterance when they assign strength

ratings to speech. Notwithstanding these advantages for rating the frequency of client speech, it is likely that including reliable ratings of the strength of client commitment language in addition to frequency would result in additional explanatory power.

Despite its strengths, this study suffers from several limitations that constrain the inferences that may be drawn from it. First, the data indicate a relatively brief and relatively small association between change talk during MI sessions and drinking outcomes. Our study demonstrated an association between a single MI session and subsequent drinking up to five weeks later. A time-limited relationship of this nature is logical if we suppose that a brief treatment like MI would confer immediate but fragile changes in motivation that must then be sustained by environmental contingencies as the treatment episode becomes more remote. This time-limited relationship between in-session events and subsequent drinking outcomes may also explain what appears to be relatively weak mediation between MICO behaviors and client change talk and drinking outcomes. Indeed, Shrout and Bolger (2002) have suggested that when predictors occur well prior to outcomes, a real effect may appear smaller or not significant.

Similarly, the percentage of variance in drinking outcomes accounted for by in-session client speech is about 3% in our study, indicating a very small association between the two. We surmise that change talk is only one of a number of in-treatment and post-treatment variables that influence drinking outcomes. Although it accounts for only a small amount of the variance in MI outcomes, client speech nevertheless retains interest for further investigation because it can potentially be influenced by therapists directly and because it functions as an immediate feedback mechanism to therapists as to how well they are using MI within any given session. Previous research from our group has indicated that the therapist's interpersonal skills, such as empathy, may account for more of a client's response during MI sessions than do the specific therapist behaviors, such as emphasizing personal choice (Moyers, Miller, & Hendrickson, 2005). It may be that how therapists implement MI behaviors, such as eliciting change talk, is as important as using them at all.

A second caveat is the nonexperimental nature of our data. The association between change talk and drinking outcomes seen in our findings could be explained by an unmeasured variable; for example, client motivation. To address this possibility we included baseline measures of client eagerness for change as a covariate in the multilevel models to predict drinking outcomes and still found an effect for change talk. Nevertheless, a stronger argument for the causal nature of change talk can only be made via experimental studies allowing it to be manipulated within treatment sessions (Spencer, Zanna, & Fong, 2005). Experimental manipulation of change talk within sessions could be demonstrated if therapists employed differential strategies to recognize, reinforce, and elicit client change talk, and if clients were randomly assigned to receive those strategies, or not, as part of their treatment. Our data, while providing the preliminary scientific foundation for further investigation of change talk as a causal mechanism, do not offer this kind of strong and direct support for change talk as an active ingredient within MI.

A third caveat in our study concerns the impact of the second MET sessions within this treatment upon drinking outcomes. Because we were not able to analyze the second session data, their impact upon changes in drinking outcomes remains unknown. Future research in this area would be strengthened by examining a series of sessions with sequential analyses and measuring the impact of that entire series upon behavioral outcomes.

Despite its limitations, this study directly addresses the knowledge gap concerning mechanisms of action for substance abuse treatments more generally (Longabaugh et al., 2005; McKay, 2007) and MI in particular (Allsop, 2007; Morgenstern & McKay, 2007). These data are

especially informative since they measure particular aspects of the therapist's behavior that are hypothesized to impact a mediating variable and then tie that mediator directly to clinical outcomes (McKay, 2007). Although we have explored change talk as a specific hypothesized ingredient within MI, it is possible that the effects of client change speech are not specific to MI but are active in other substance abuse treatments as well. We have reason to believe this might be so, since change talk is also predictive of outcome in cognitive-behavior therapy and twelve-step facilitation treatments for problem drinking (Moyers et al., 2007). Further, client commitments, whether written (Hall, Havassy, & Wasserman, 1990), spoken (Mahrer, Gagnon, Fairweather, Boulet, & Herring, 1994), or both (Putnam, Finney, Barkley, & Bonner, 1994), are predictive of outcomes across target behaviors and treatment modalities. While the importance of client speech may not be specific to MI (e.g., Russell, 1998), the identification of change talk as a construct, as well as the explicit strategies for selective attention to it by the therapist, appear to be unique among therapeutic approaches and worthy of future investigation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The research discussed in this paper was funded entirely by NIAAA R01 AA13696-01

References

Aharonovich E, Amrhein PC, Bisaga A, Nunes EV, Hasin DS. Cognition, commitment language, and behavioral change among cocaine-dependent patients. Psychology of Addictive Behaviors 2008;22 (4):557–562. [PubMed: 19071981]

Allsop S. What is this thing called motivational interviewing? Addiction 2007;102:343–345. [PubMed: 17298637]

Amrhein PC. How does motivational interviewing work? What client talk reveals. Journal of Cognitive Psychotherapy: An International Quarterly 2004;18(4):323–336.

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

Baer JS, Beadnell B, Garrett SB, Hartzler B, Wells EA, Peterson PL. Adolescent change language within a Brief Motivational Intervention and substance use outcomes. Psychology of Addictive Behaviors 2008;22(4):570–575. [PubMed: 19071983]

Bakeman, R.; Gottman, JM. Observing interaction. 2. New York: Cambridge University Press; 1997.

Bakeman, R.; Quera, V. Analyzing interaction: Sequential analysis with SDIS and GSEQ. New York: University of Cambridge Press; 1995.

Bakeman R, Quera V, McArthur D, Robinson BF. Detecting sequential patterns and determining their reliability with fallible observers. Psychological Methods 1997;2(4):357–370.

Bandura A. Human agency in social cognitive theory. American Psychologist 1989;44:1175–1184. [PubMed: 2782727]

Baron RM, Kenney DA. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. Journal of Personality and Social Psychology 1986;51(6):1173–1182. [PubMed: 3806354]

Carbonari, JP.; DiClemente, CC.; Zweben, A. A Readiness to Change Scale: Its development, validation, and usefulness. Paper presented at the annual meeting of the Association for Advancement of Behavior Therapy.; 1994.

Carroll KM, Connors GJ, Cooney NL, DiClemente CC, Donovan DM, Kadden RR, et al. Internal Validity of Project MATCH Treatments: Discriminability and Integrity. Journal of Consulting and Clinical Psychology 1998;66(2):290–303. [PubMed: 9583332]

Carroll KM, Ball SA, Nich C, Martino S, Frankforter TL, Farentinos C, et al. Motivational interviewing to improve treatment engagement and outcome in individuals seeking treatment for substance abuse: A multisite effectiveness study. Drug and Alcohol Dependence 2006;81(3):301–312. [PubMed: 16169159]

Catley D, Harris KJ, Mayo MS, Hall S, Okuyemi KS, Boardman T, et al. Adherence to principles of motivational interviewing and client within-session behavior. Behavioural and Cognitive Psychotherapy 2006;34:43–56.

Cicchetti DV. Guidelines, criteria and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychological Assessment 1994;6:284–290.

Cohen J. A coefficient of agreement for nominal scales. Educational and Psychological Measurement 1960;20:37–46.

Connors GJ, Tonigan JS, Miller WR. A longitudinal model of intake symptomatology, AA participation and outcome: Retrospective study of the Project MATCH outpatient and aftercare samples. Journal of Studies on Alcohol 2001;62(6):817–825. [PubMed: 11838919]

DiClemente CC, Carbonari JP, Montgomery RPG, Hughes SO. Alcohol abstinence self-efficacy scale. Journal of Studies on Alcohol 1994;55(2):141–148. [PubMed: 8189734]

DiClemente CC, Hughes SO. Stages of change profiles in outpatient alcoholism treatment. Journal of Substance Abuse 1990;2:217–235. [PubMed: 2136111]

Gardner, W. On the reliability of sequential data: Measurement, meaning, and correction. In: Gottman, JM., editor. The Analysis of Change. Hillsdale, N.J: Erlbaum; 1995. p. 339-359.

Gaume J, Gmel G, Daeppen JB. Brief alcohol interventions: Do counsellors' and patients' communication characteristics predict change? Alcohol & Alcoholism 2008;43:62–69. [PubMed: 17942439]

Gaume J, Gmel G, Faouzi M, Daeppen JB. Counsellor behaviours and patient language during brief motivational interventions: A sequential analysis of speech. Addiction 2008;103(11):1793–1800. [PubMed: 19032529]

Hall SM, Havassy BE, Wasserman DA. Commitment to abstinence and acute stress in relapse to alcohol, opiates, and nicotine. Journal of Consulting and Clinical Psychology 1990;58(2):175–181. [PubMed: 2335634]

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

Houck, JM.; Moyers, TB. What you do matters: Therapist influence on client behavior during motivational interviewing sessions. Paper presented at the First International Addiction Summit; Melbourne, Australia. 2008.

Hox, J. Multilevel analyses: Techniques and applications. Mahwah, NJ: Lawrence Erlbaum Associates; 2000.

Kreft, I.; de Leeuw, J. Introducing multilevel modeling. Thousand Oaks, CA: Sage Publications; 1998.

Krull JL, MacKinnon DP. Multilevel modeling of individual and group level mediated effects. Multivariate Behavioral Research 2001;36(2):249–277.

Longabaugh R, Donovan DM, Karno MP, McCrady BS, Morgenstern J, Tonigan JS. Active ingredients: How and why evidence-based alcohol behavioral treatment interventions work. Alcoholism: Clinical and Experimental Research 2005;29(2):235–247.

MacKinnon, DP. Introduction to statistical mediation analysis. Mahwah, NJ: Erlbaum Psych Press; 2008.

MacKinnon DP, Lockwood CM, Hoffman JM, West SG, Sheets V. A comparison of methods to test mediation and other intervening variable effects. Psychological Methods 2002;7(1):83–104. [PubMed: 11928892]

MacKinnon DP, Fritz MS, Williams J, Lockwood CM. Distribution of the product confidence limits for the indirect effect: Program PRODCLIN. Behavior Research Methods 2007;39(3):384–389. [PubMed: 17958149]

Mahrer AR, Gagnon R, Fairweather DR, Boulet DB, Herring CB. Client commitment and resolve to carry out postsession behaviors. Journal of Counseling Psychology 1994;41(3):407–414.

Martin, T.; Moyers, TB.; Houck, JM.; Christopher, PJ.; Miller, WR. Motivational Interviewing Sequential Code for Observing Process Exchanges (MI-SCOPE) coder's manual. 2005. Retrieved (April 15, 2009) from http://casaa.unm.edu/download/scope.pdf

McConnaughy EA, Prochaska JO, Velicer WF. Stages of change in psychotherapy: Measurement and sample profiles. Psychotherapy: Theory, Research, and Practice 1983;20:368–375.

McKay JR. Lessons learned from psychotherapy research. Alcoholism: Clinical and Experimental Research 2007;31(S3):48–54.

Miller WR. Motivational Interviewing with problem drinkers. Behavioural Psychotherapy 1983;11:147–172.

Miller, WR. The Motivational Interviewing Skills Code (MISC) manual. 2000. Retrieved (April 15, 2009) from http://casaa.unm.edu/download/misc1.pdf

Miller WR, Del Boca FK. Measurement of drinking behavior using the Form 90 family of instruments. Journal of Studies on Alcohol 1994;S12:112–118.

Miller WR, Moyers TB. Eight stages in learning motivational interviewing. Journal of Teaching in the Addictions 2007;5(1):3–17.

Miller WR, Moyers TB, Amrhein PC, Rollnick S. A consensus statement on defining change talk. MINT Bulletin 2006;13(2):6–7.

Miller, WR.; Rollnick, S. Motivational Interviewing: Preparing people to change addictive behaviors. New York: Guilford Press; 1991.

Miller, WR.; Rollnick, S. Motivational Interviewing: Preparing people to change. 2. New York: Guilford Press; 2002.

Miller WR, Yahne CE, Moyers TB, Martinez J, Pirritano M. A randomized trial of methods to help clinicians learn motivational interviewing. Journal of Consulting and Clinical Psychology 2004;72:1050–1062. [PubMed: 15612851]

Morgenstern J, McKay JR. Rethinking the paradigms that inform behavioral treatment research for substance use disorders. Addiction 2007;102(9):1377–1389. [PubMed: 17610541]

Moyers TB, Martin T. Therapist influence on client language during motivational interviewing sessions. Journal of Substance Abuse Treatment 2006;30:245–251. [PubMed: 16616169]

Moyers TB, Martin T, Christopher PJ, Houck JM, Tonigan JS, Amrhein PC. Client language as a mediator of motivational interviewing efficacy: Where is the evidence? Alcoholism: Clinical and Experimental Research 2007;31(S3):40–47.

Moyers TB, Miller WR, Hendrickson SML. How does motivational interviewing work? Therapist interpersonal skill predicts client involvement within motivational interviewing sessions. Journal of Consulting and Clinical Psychology 2005;73(4):590–598. [PubMed: 16173846]

Preacher KJ, Hayes AF. Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. Behavior Research Methods 2008;40(3):879–891. [PubMed: 18697684]

Project MATCH Research Group. Matching alcoholism treatments to client heterogeneity: Project MATCH posttreatment drinking outcomes. Journal of Studies on Alcohol 1997;58:7–29. [PubMed: 8979210]

Project MATCH Research Group. Therapist effects in three treatments for alcohol problems. Psychotherapy Research 1998;8(4):455–474.

Putnam DE, Finney JW, Barkley PL, Bonner MJ. Enhancing commitment improves adherence to a medical regimen. Journal of Consulting and Clinical Psychology 1994;62(1):191–194. [PubMed: 8034823]

Raudenbush, SW.; Bryk, AS. Hierarchical linear models: Applications and data analysis methods. Thousand Oaks, CA: Sage Publications; 2002.

Raudenbush, SW.; Bryk, AS.; Congdon, R. HLM 6: Hierarchical linear modeling and nonlinear modeling. Chicago: Scientific Software International; 2004.

Russell RL. Linguistic psychotherapy research: New directions and promising findings. Journal of Clinical Child Psychology 1998;27(1):17–27. [PubMed: 9561933]

Rychtarik, RG.; Miller, WR.; Tonigan, JS. The alcohol involvement matching hypothesis and findings. In: Longabaugh, RH.; Wirtz, PW., editors. Project MATCH hypotheses: Results and causal chain analyses. Vol. 8. Rockville, MD: NIAAA; 2001. p. 44-60.Project MATCH Monograph Series

Shrout PE, Bolger N. Mediation in experimental and nonexperimental studies: New procedures and recommendations. Psychological Methods 2002;7(4):422–445. [PubMed: 12530702]

Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. Psychological Bulletin 1979;86(2):420–428. [PubMed: 18839484]

Skinner HA, Allen BA. Differential assessment of alcoholism -- Evaluation of the Alcohol Use Inventory. Journal of Studies on Alcohol 1983;44:852–862. [PubMed: 6645545]

Sobel. Asymptotic confidence intervals for indirect effects in structural equation models. Sociological Methodology 1982;13:290–312.

Spencer SJ, Zanna MP, Fong GT. Establishing a causal chain: Why experiments are often more effective than mediational analyses in examining psychological processes. Journal of Personality and Social Psychology 2005;89(6):845–851. [PubMed: 16393019]

Strang J, McCambridge J. Can the practitioner correctly predict outcome in motivational interviewing? Journal of Substance Abuse Treatment 2004;27(1):83. [PubMed: 15223098]

Taylor AB, MacKinnon DP, Tein JY. Tests of the three-path mediated effect. Organizational Research Methods 2008;11(2):241–269.

Tonigan JS, Miller WR, Brown JM. The reliability of Form 90: An instrument for assessing alcohol treatment outcome. Journal of Studies on Alcohol 1997;58:358–364. [PubMed: 9203116]

Wanberg KW, Horn JL, Foster FM. A differential assessment model for alcoholism: The scales of the Alcohol Use Inventory. Journal of Studies on Alcohol 1975;38(3):512–543. [PubMed: 859335]

Wickens, TD. Models for behavior: Stochastic processes in psychology. San Francisco: W.H. Freeman & Company; 1982.

Williams J, MacKinnon DP. Resampling and Distribution of the Product Methods for Testing Indirect Effects in Complex Models. Structural Equation Modeling 2008;15:23–51. [PubMed: 20179778]

Witte, G.; Wilber, CH. A case study in clinical supervision: Experience from Project MATCH. In: Carroll, KM., editor. Improving Compliance with Alcoholism Treatment. Vol. 6. Rockville, MD: National Institute on Alcohol Abuse and Alcoholism; 1997. p. 73-86.Project MATCH Monograph Series

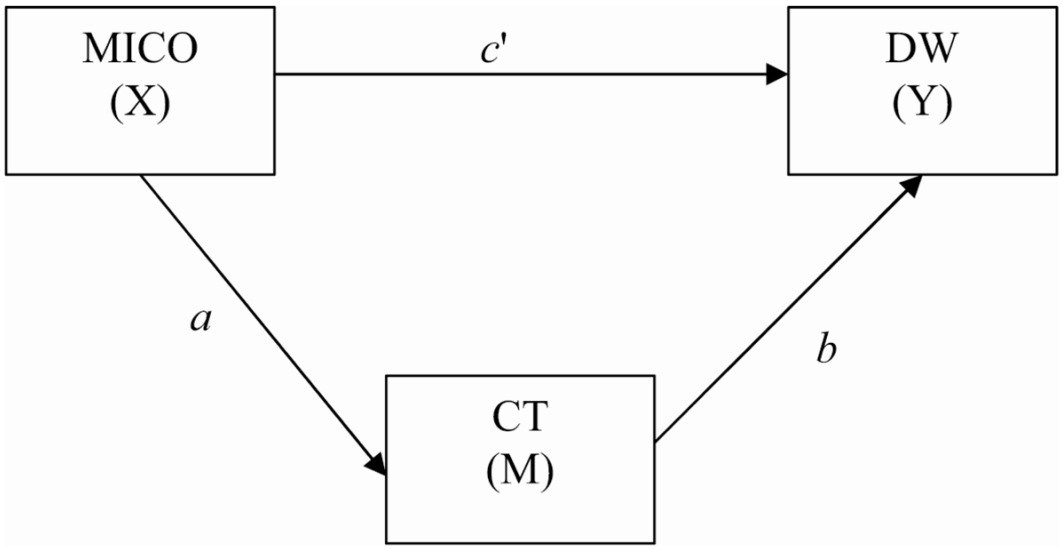

Figure 1.
Mediation model of MI-consistent behaviors (MICO), change talk (CT), and Drinks per week (DW).

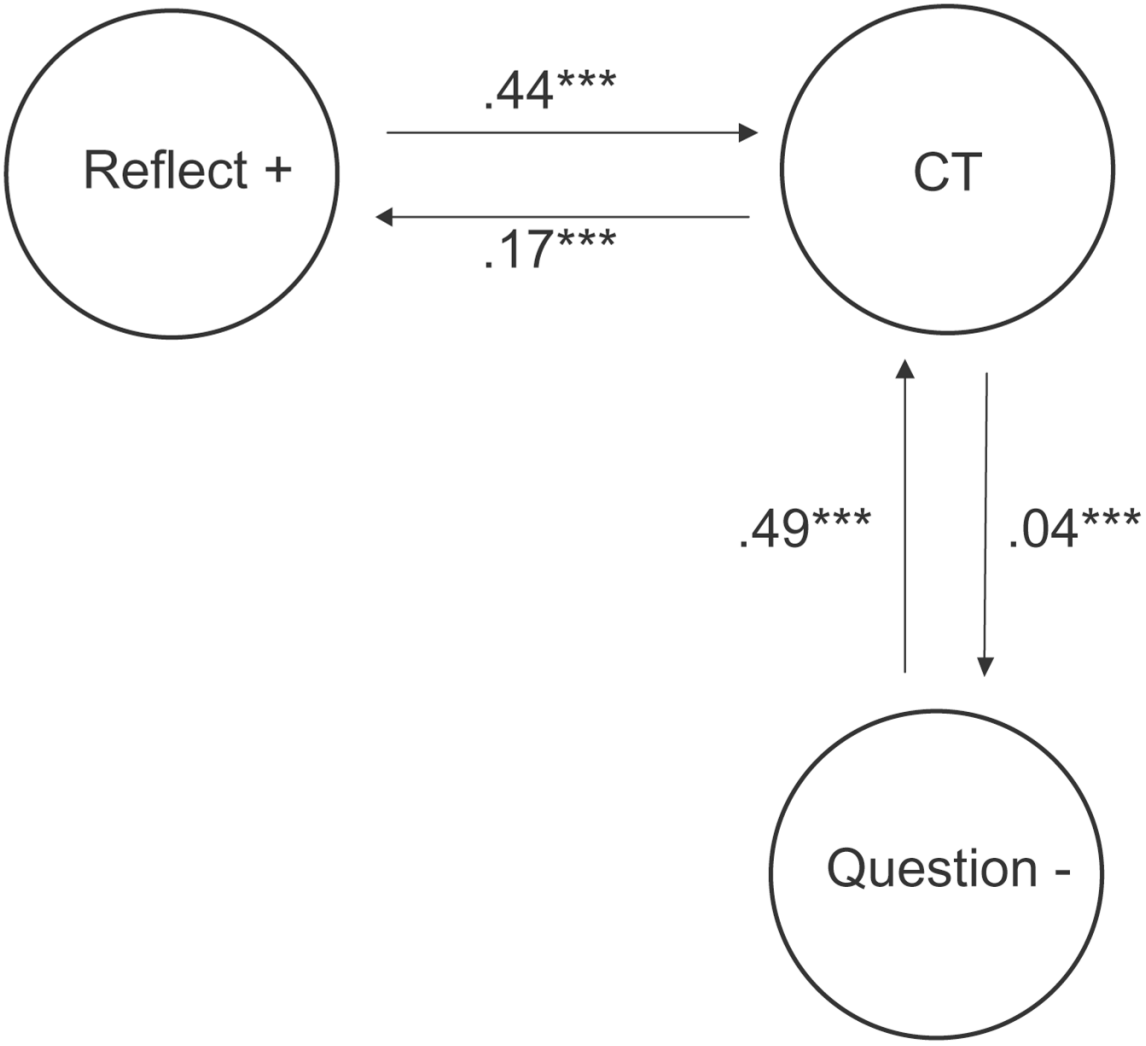

Figure 2.
State transition diagram for change talk (CT), reflections of change talk (Reflect +), and questioning the negative aspects of the target behavior (Question −). All transitions shown here were more probable than expected by chance, $p < .001$.

Table 1

Descriptive Statistics on Outcome Measures and Covariates

	Drinks per week						AIM	Self- efficacy	Readiness
	Baseline	**Week 1**	**Week 2**	**Week 3**	**Week 4**	**Week 5**			
M	76.3	14.8	12.3	8.0	6.1	6.8	30.1	60.2	10.8
SD	78.8	50.1	39.6	27.5	18.7	21.8	12.0	18.7	1.7
N	118	112	117	117	118	118	106	103	116

Note. AIM = third-order measure of Alcohol Involvement from the Alcohol Use Inventory, measured at intake; Self-efficacy = Alcohol Abstinence Self-Efficacy scale, measured at intake; Readiness = measure of readiness derived from the University of Rhode Island Change Assessment scale by adding the Contemplation, Action, and Maintenance subscale scores and subtracting the Precontemplation subscale score, measured at intake.

Table 2

ICCs for Summary Variables

Summary variable	*f*	Min	Max	*M*	*SD*	*ICC*
CT	6475	3	118	54.87	24.34	.88
CCT	2202	1	75	18.66	14.08	.86
FN/Ask	17673	31	307	149.77	60.17	.96
MICO	2168	3	52	18.37	10.35	.84
MIIN	1025	0	97	8.69	11.18	.79
Question	7904	11	156	66.98	31.43	.98
Question (positive)	286	0	15	2.42	2.88	.75
Question (negative)	837	0	29	7.09	5.91	.60
Closed Question	6130	3	125	51.95	27.09	.98
Open Question	1774	2	51	15.03	8.53	.90
Reflect	7207	5	120	61.08	26.28	.92
Reflect CT	2315	0	64	19.62	12.38	.85
Reflect CCT	547	0	25	4.64	4.81	.73
Reflect Other	92	0	7	0.78	1.31	.56
Simple Reflection	5292	4	92	44.85	20.37	.72
Complex Reflection	1915	0	63	16.23	11.47	.49

Note. f, Min, Max, *M*, & *SD* computed on full sample of 118 tapes. *ICC* = single measures intraclass correlation, computed for the three primary coders on a reliability sample of 40 tapes. CT = Change talk, CCT = Counterchange talk, FN = Follow/neutral, MICO = MI-Consistent behaviors, MIIN = MI-Inconsistent behaviors.

Table 3

Regressions of client speech on therapist behaviors; mediation path a

Variable	*B*	*SE B*	95% CI	β	*t*
Change talk					
MICO	0.30	0.06	0.19, 0.42	0.47	5.33***
MIIN	−0.20	0.19	−0.57, 0.18	−0.09	−1.02
Counter-change talk					
MICO	0.10	0.03	0.03, 0.17	0.27	2.99**
MIIN	0.29	0.11	0.07, 0.52	0.23	2.57**

Note. Change talk adjusted R^2 = .036, $p < .05$; Counter-change talk adjusted R^2 = .096, $p < .01$; Percent change talk R^2 = .031, $p = .06$; MICO = Motivational Interviewing Consistent behaviors; MIIN = Motivational interviewing inconsistent behaviors

* $p < .05$

** $p < .01$

*** $p < .001$

Table 4

Multilevel model of the effects of change talk on drinks per week; mediation path b

Fixed effect	Coeff	*SE*	Event rate ratio	95% CI	*t* ratio
Intercept of DW					
READI	−0.434	0.009	0.648	0.637, 0.659	−50.188
CT	−0.024	0.001	0.976	0.995, 0.977	−34.548
Linear slope of DW					
READI	0.050	0.008	1.051	1.034, 1.068	5.987
CT	0.005	0.001	1.005	1.003, 1.006	6.616
Quadratic slope of DW					
READI	0.054	0.003	1.055	1.048, 1.062	15.373
CT	0.004	0.000	1.004	1.003, 1.004	13.871

Note. All *p*'s < .01. Population-average model with robust standard errors. DW = Drinks per Week; READI = Readiness to change as measured by the University of Rhode Island Change Assessment scale; CT = Change talk

Table 5

Multilevel model of the effects of MICO on drinks per week; mediation path c

Fixed effect	Coeff	*SE*	Event rate ratio	95% CI	*t* ratio
Intercept of DW					
READI	−0.48	0.09	0.62	0.61, 0.63	−57.84
MICO	−0.00	0.00	0.99	0.99, 1.00	−8.59
Linear slope of DW					
READI	0.06	0.01	1.06	1.04, 1.08	7.21
MICO	−0.00	0.01	1.00	1.00, 1.00	−0.74^
Quadratic slope of DW					
READI	0.06	0.00	1.06	1.05, 1.07	18.49
MICO	0.00	0.00	1.00	1.00, 1.00	2.87

Note. Except where noted, all *p*'s < .01; ^ = *p* > .50. Population -average model with robust standard errors. DW = Drinks per Week; READI = Readiness to change as measured by the University of Rhode Island Change Assessment scale; MICO = Motivational Interviewing Consistent behaviors.

Table 6

Multilevel model of the effects of change talk and MICO on drinks per week; mediation path c′

Fixed effect	Coeff	*SE*	Event rate ratio	95% CI	*t* ratio
Intercept of DW					
READI	−0.44	0.01	0.65	0.64, 0.66	−51.76
CT	−0.03	0.00	0.97	0.97, 0.98	−36.96
MICO	0.00	0.00	1.00	1.00, 1.00	5.04
Linear slope of DW					
READI	0.05	0.01	1.05	1.04,1.07	6.25
CT	0.01	0.00	1.01	1.00, 1.01	8.68
MICO	0.00	0.00	1.00	1.00, 1.00	−3.95
Quadratic slope of DW					
READI	0.05	0.00	1.06	1.05, 1.06	16.31
CT	0.00	0.00	1.00	1.00, 1.00	13.76
MICO	0.00	0.00	1.00	1.00, 1.00	−2.40

Note. All *p*'s < .01. Population -average model with robust standard errors. DW = Drinks per Week; READI = Readiness to change as measured by the University of Rhode Island Change Assessment scale; CT = Change talk; MICO = Motivational Interviewing Consistent behaviors.

Table 7

Transition Matrix for SCOPE Behaviors

	Subsequent behavior											
Initial behavior	**CT**	**CCT**	**FN/ASK**	**sMICO**	**MIIN**	**QPOS**	**QNEG**	**QNEUT**	**RCT**	**RCCT**	**REF**	**OTHER**
CT	$.24^{+++}$	$.06^{+++}$	.16	$.05^{+++}$	.01	$.01^{+++}$	$.04^{+++}$	.01	$.17^{+++}$	.01	.05	.11
CCT	$.15^{+++}$	$.15^{+++}$	.17	$.03^{-}$	.01	$.02^{+++}$	.02	.11	.04	$.12^{+++}$	.08	.10
FN/ASK	.06	.02	.08	$.05^{+++}$	$.02^{++}$	.01	$.02^{++}$	$.21^{+++}$	.04	.01	$.18^{+++}$	$.30^{+++}$
sMICO	.12	.02	$.36^{+++}$	$.06^{+++}$	$.03^{+++}$	$.00^{-}$	.01	.08	.04	$.00^{--}$	.03	.25
MIIN	.09–	.05	.28	$.06^{+++}$	$.08^{+++}$	.00	$.00^{-}$	.07	.01	.00	.03	$.32^{+++}$
QPOS	$.26^{+++}$	$.36^{+++}$	.32	$.00^{--}$	.01	.01	$.00^{-}$	.00	$.01^{--}$	.01	.01	.02
QNEG	$.49^{+++}$	$.06^{+++}$	$.40^{+++}$	.00	.00	$.00^{-}$	$.00^{--}$	.00	$.02^{--}$	$.00^{--}$	.00	.02
QNEUT	.10	.03	$.81^{+++}$	.00	.00	.00	.00	.01	.00	.00	.00	.02
RCT	$.44^{+++}$	.03	$.26^{--}$	$.03^{--}$	.01	$.01^{++}$	$.03^{+++}$	.05	$.03^{-}$	.00	.01	.10
RCCT	$.07^{-}$	$.38^{+++}$	.29	$.01^{--}$	.02	$.03^{+++}$	.02	.06	$.01^{-}$	.01	.01	.09
REF	.06	.03	$.62^{+++}$	.02	.01	.00	.01	$.10^{-}$	.01	.00	.02	.11
OTHER	.04	.01	$.29^{-}$	.03	.02	.00	.01	.08	.01	.00	.02	$.50^{+++}$

Note. These transition probabilities give the probability that if the row category occurs, the next category of speech will be the column category. For example, given that the client emits CT, the first row, fourth column cell C_{14} (.05) gives the probability that the next utterance will be a therapist utterance falling into the MICO category. CT = Change Talk, CCT = Counterchange Talk, FN/ASK = Follow/Neutral + Ask, sMICO = sequential MI-consistent, MIIN = MI-Inconsistent, QPOS = Question Positive, QNEG = Question Negative, QNEUT = Question Neutral, RCT = Reflect CT, RCCT = Reflect CCT, REF = Reflect Neutral + Reflect Both, Other = Filler, Feedback, Opinion, Raise Concern, Structure.

$^{+}$ More probable than expected by chance, $p < .05$

$^{++}$ More probable than expected by chance, $p < .01$,

$^{+++}$ more probable than expected by chance, $p < .001$,

$^{-}$ Less probable than expected by chance, $p < .05$,

$^{--}$ Less probable than expected by chance, $p < .01$,

$^{---}$ Less probable than expected by chance, $p < .001$.