2-13-2014

RELATING CLIENT CHANGE LANGUAGE AND SAFER-SEX OUTCOMES IN A GROUP-DELIVERED MOTIVATIONAL ENHANCEMENT THERAPY (GMET) INTERVENTION FOR DETAINED ADOLESCENTS

Lisa Glynn

Lisa H. Glynn
Candidate

Psychology
Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Theresa Moyers, Ph.D., Chairperson

Barbara McCrady, Ph.D.

Angela Bryan, Ph.D.

James Selig, Ph.D.

RELATING CLIENT CHANGE LANGUAGE AND SAFER-SEX OUTCOMES IN A GROUP-DELIVERED MOTIVATIONAL ENHANCEMENT THERAPY (GMET) INTERVENTION FOR DETAINED ADOLESCENTS

By

LISA H. GLYNN

B.S., Psychology, University of Washington, 2001
M.S., Psychology, University of New Mexico, 2009

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
Psychology

The University of New Mexico
Albuquerque, New Mexico

December 2013

DEDICATION

To Cocoa, Purdy, and Snowball, who all deserved dogtorates.

ACKNOWLEDGMENTS

I gratefully acknowledge the support of all who helped me complete this dissertation. I extend my utmost gratitude to my mentor and committee chair, Dr. Theresa Moyers, who guided me from proposal to defense, and reminded me constantly what is Real. I sincerely thank my committee members, Drs. Barbara McCrady, Angela Bryan, and James Selig, for their invaluable expertise for this project and their years of investment in my academic growth. I also owe great appreciation to my funders, Dr. McCrady, Dr. Moyers, NIAAA training grant T32AA018108, and the Department of Psychology. I also am indebted to my coders, my lab mates, and Project MARS participants, interventionists, and research staff. I also wish to thank my husband and partner, Michael, and our family, friends, neighbors, and Albuquerque's jazz community, for their perpetual love and understanding throughout the past seven years.

RELATING CLIENT CHANGE LANGUAGE AND SAFER-SEX OUTCOMES IN A GROUP-DELIVERED MOTIVATIONAL ENHANCEMENT THERAPY (GMET) INTERVENTION FOR DETAINED ADOLESCENTS

By

Lisa H. Glynn

B.S., Psychology, University of Washington, 2001
M.S., Psychology, University of New Mexico, 2009
Ph.D., Psychology, University of New Mexico, 2013

ABSTRACT

"Change talk" refers to client statements in favor of changing a particular target behavior. In individual motivational interviewing (MI) psychotherapy sessions, change talk is a predictor of eventual change for a variety of target behaviors. Growing evidence points to change talk as a potential causal mechanism of that treatment, but it remains unclear whether it might underlie the efficacy of group-delivered MI as well. Group-delivered MI and its adaptations have been employed widely across treatment settings and target behaviors; however, such groups are relatively unstudied, they have varied considerably in their elements and delivery, and their outcomes have shown mixed results. Identifying and implementing the active ingredients of group-delivered MI interventions could increase treatment efficiency. This study sought to understand whether change talk, in combination with other theoretically relevant variables, might explain the efficacy of a group Motivational Enhancement Therapy (gMET) intervention known to be associated with reductions in risky sex behaviors. A secondary analysis of audio-recordings from Project MARS captured the in-session language of a diverse

sample of 200 detained adolescents within 58 gMET interventions, which was reduced to 45 participants within 27 groups after exclusions. Recordings were sequentially coded for change talk and other client and clinician behaviors using an adaptation of the MISC 2.5 coding system and CACTI software. Multiple regression analyses used group-level change talk, client sex, and baseline questionnaire scores to predict clients' individual and group-averaged safer-sex behaviors at three-month follow-up. Safer-sex behaviors were measured by two composite variables of items from the Sexual History Questionnaire (Risky Sex Index and Safer Sex). Three of four full models were statistically significant, but only baseline questionnaire scores were significant predictors of those three-month scores and not change talk or client sex. Possible interpretations are offered, and future directions are discussed.

Keywords: group, safer sex, change talk, motivational interviewing, Motivational Enhancement Therapy, adolescent, incarcerated

Table of Contents

List of Figures

List of Tables

Introduction

Group-delivered motivational interviewing (MI) is being implemented widely for the treatment of various disorders in adolescents and adults. However, the literature has shown limited and contradictory evidence of efficacy for MI groups, and little concrete knowledge or even unified theory exists to explain how and why group treatment might work. This review summarizes the relevant findings about group treatment, MI and its adaptations, and the hypothesized mechanisms of MI that might be responsible for change.

Why Group Therapy?

Group-delivered therapy is implemented instead of individual therapy for a number of reasons. For example, (1) it is an efficient use of therapist resources and is cost-effective (Center for Substance Abuse Treatment [CSAT], 2005; French et al., 2008; Kaminer, Burleson, & Goldberger, 2002), (2) it produces outcomes comparable to or even better than those of individual therapy (Burleson, Kaminer, & Dennis, 2006; Renjilian et al., 2001; Weiss, 2004), (3) it provides opportunities for clients to build interpersonal skills and support, and to gain knowledge or information (CSAT, 2005; Weiss, 2004; Yalom, 1985), and (4) it offers therapeutic benefits beyond the content of the intervention, including hope, universality, and altruism (Yalom, 1985). Thus, ideally the meeting of group members is not only an efficient way to deliver treatment or a forum for learning and practicing skills, but also a shared interpersonal experience—a uniquely group-created, sum-greater-than-its-parts milieu—that is hypothesized to set the conditions that foster change.

Groups and adolescents. There is good evidence of efficacy for group-delivered therapy with adolescents (Hoag & Burlingame, 1997; Waldron & Turner, 2008). It also is considered a developmentally appropriate intervention for adolescents, who tend to rely heavily upon peer support when considering difficult changes (O'Leary, et al., 2002; Stern, Meredith, Gholson, Gore, & D'Amico, 2007). However, grouping adolescents is sometimes associated with worse outcomes, because the dysfunctional or antisocial behaviors of a few group members can be "contagious" (e.g., Dishion & Dodge, 2005; Kaminer, 2005). Other studies have contradicted this finding, showing that "contagion" should not be a concern, and that more deviant adolescents actually might benefit from sharing a group setting with peers with less severe conduct problems (Burleson, Kaminer, & Dennis, 2006). Thus, the cumulative evidence for group therapy in adolescents is mixed.

Motivational Interviewing

Motivational interviewing (MI) is a "client-centered, directive method for enhancing intrinsic motivation to change by exploring and resolving ambivalence" (Miller & Rollnick, 2002, p. 25). It is an evidence-based therapeutic style that has shown efficacy for an array of target health behaviors (CSAT, 2006; Hettema, Steele, & Miller, 2005), including alcohol (e.g., Vasilaki, Hosier, & Cox, 2006), marijuana (e.g., Lundahl, Kunz, Brownell, Tollefson, & Burke, 2010; Walker et al., 2006), and safer sex (Lundahl et al., 2010). Even brief MI-based interventions have been effective in producing long-lasting behavioral change in clients (e.g., Burke, Arkowitz, & Menchola, 2003; Dunn, Deroo, & Rivara, 2001; Hettema, Steele, & Miller, 2005), even as long as three years post-intervention (Project MATCH Research Group, 1998; Schermer, Moyers, Miller, &

Bloomfield, 2006), which makes them particularly attractive for maximizing the number of clients served. Despite the balance of evidence showing that MI is efficacious, the mechanisms by which MI operates remain speculative.

In search of mechanisms. Establishing the efficacy, effectiveness, and efficiency of a treatment such as MI is an important first step. However, this shows only that a treatment works, without explaining how or why. In order to fully understand why and how MI operates, it will be important to examine its mechanisms, or "…those processes or events that lead to and cause therapeutic change" (Kazdin & Nock, 2003, p. 1117). This research is underway in MI, but few solid conclusions have been reached.

It is hypothesized that causal clinician behaviors in MI include a combination of "relational" and "technical" factors (Miller & Rose, 2009). Among these are common therapeutic factors (e.g., accurate empathy, acceptance, collaboration, support of client autonomy, and evocation), along with specific factors (e.g., attending to client "change talk") that are emphasized exclusively in MI. These are discussed below.

Relational factors. Most of the relational factors of MI hearken back to those employed by Rogers in his client-centered psychotherapy (Rogers, 1965). The primary hypothesized relational variables of MI are accurate empathy and MI "spirit" (Miller & Rollnick, 2002), but relatively little concrete evidence exists to support either element as a potential mechanism of MI (e.g., Apodaca & Longabaugh, 2009). For that reason, this review will be brief, and focused primarily upon the theoretical underpinnings for the global ratings that are included in MI-based coding systems.

Empathy. "Expressing empathy" (Miller & Rollnick, 2002, p. 7) is considered one of four general principles of MI. Its meaning is most closely related to *understanding*

the client, and employing it in MI typically involves the use of accurate and complex reflections of client speech (e.g., Houck, Moyers, Miller, Glynn, & Hallgren, 2010).

Support for accurate empathy is limited but growing. A meta-analysis found mixed and insufficient support for empathy alone as a potential mechanism of MI, showing that empathy did not distinguish MI from other therapies (Apodaca & Longabaugh, 2009). Since then, another study has linked clinician empathy ratings to client outcomes within a group-delivered intervention for adolescent marijuana users (Engle, Macgowen, Wagner, & Amrhein, 2010); because the intervention was MI-based, the question of whether empathy or another element of MI was responsible for outcomes remains unresolved. More recently, a review of clinician empathy in MI concluded that it accounts for significant variance in client outcomes and therapeutic alliance, and should be emphasized in substance-abuse treatment to prevent harm to clients (Moyers & Miller, 2012).

MI spirit. MI spirit refers to a combination of collaboration, evocation, and support for client autonomy (e.g., Moyers, Martin, Manuel, & Hendrickson, 2005). Collaboration refers to *working with* clients to help them choose their own goals, evocation refers to drawing out their reasons for change, and autonomy support means that clients should have choice in whether and how to change (e.g., Houck et al., 2010).

As with empathy, a meta-analysis failed to find support for MI spirit as a determinant of client outcomes in MI (Apodaca & Longabaugh, 2009). Just one more recent study has found support for MI spirit: In an investigation of marijuana cessation in teens, MI spirit predicted three-month outcomes (McCambridge, Day, Thomas, & Strang, 2011).

Acceptance. Acceptance refers to *unconditional positive regard* toward clients, which again draws from Rogers and is believed to paradoxically free a client to change (Miller & Rollnick, 2002, p. 7). No research evidence is available to support acceptance as a causal element in MI.

Direction. As suggested by the definition of MI offered earlier, MI involves clinician use of direction, or *focus* upon a specific target behavior. This breaks from classic Rogerian therapy in that MI is intentionally directional (Miller & Rollnick, 2002), and a skillful MI session notably lacks the wandering quality found in non-directional therapies (Martin et al., 2008). Similarly, no research evidence supports this hypothesized causal element.

Self-Exploration. Self-exploration rates the client on exploration of "personally relevant material" (Houck et al., 2010, p. 11). This construct was explored originally by Truax and Carkhuff (1967), who, like Rogers, studied client-centered psychotherapy. Some research has linked client self-exploration to clinician relational variables such as accurate empathy, respect, genuineness, and warmth (e.g., Hountras & Anderson, 1969; Kurtz & Grummon, 1972; Truax, 1968), but this relation has not been studied extensively within MI.

Technical factor: Clinicians' differential attention to change talk. A unique technical factor of MI is clinicians' differential attention to two types of client language—change talk and sustain talk—through selective evocation and consequation (e.g., Miller & Rose, 2009). Change talk refers to self-motivational client speech in favor of change, and its opposite, sustain talk, denotes client speech against change or in support of the status quo (Miller & Rollnick, 2002). Both types of language have been

used in research by dividing them into several subcategories (i.e., desire, ability, reasons, need, commitment, taking steps, and other; e.g., Houck et al., 2010), examining their frequencies or relative proportions of change and sustain talk, or tracking their strength through individually parsed units of client speech (e.g, Amrhein, Miller, Yahne, Palmer, & Fulcher, 2003). Those conceptualizations of client language have facilitated testing of a model that describes the role of this technical factor within MI.

A proposed causal chain for change talk. One theory concerning the mechanisms of MI has proposed a causal chain that links particular clinician behaviors, client change talk, and client outcomes (e.g., Miller & Rose, 2009; Moyers & Martin, 2006). Several studies have found support for the three paths of this mediational model.

Clinician behaviors predict client outcomes. Much evidence has shown that MI interventions are efficacious, which suggests that either specific content components of MI interventions, their delivery by clinicians, or an interaction of the two are responsible for client outcomes. The causal chain for change talk hypothesizes that when clinicians employ MI-consistent behaviors (i.e., affirmation, emphasizing client control, providing support, and asking permission before providing advice or raising concerns) and avoid MI-inconsistent behaviors (i.e., confronting, warning, directing, and providing advice or raising concerns without permission), their clients tend to experience improved outcomes. One study related therapist MI-consistent behaviors with reduced drinks per week in a client sample from Project MATCH (Moyers et al., 2009). Another found that a clinician MI-inconsistent behavior, confrontation, across conditions was associated with worse client drinking outcomes a year later (Miller, Benefield, & Tonigan, 1993). Although the

direct evidence for MI-consistent clinician behaviors predicting improved client outcomes is limited, there is some support for this link of the causal chain.

Change talk predicts client change. Change talk has been assigned primary importance within MI because it predicts, and is hypothesized to cause, favorable client outcomes (e.g., Moyers & Martin, 2006). Growing empirical evidence shows that change talk signals subsequent positive behavior change in clients, and conversely, that sustain talk predicts worse outcomes for clients. For example, an evaluation of client language in Project MATCH showed that the frequency of change talk predicted later client drinking outcomes (Moyers et al., 2007; Moyers et al., 2009). Likewise, a study of college drinkers related change-talk frequency to reduced drinking at three months in clients who received MI plus feedback (Vader, Walters, Prabhu, Houck, & Field, 2010). Subcategories of change and sustain talk also have been shown to predict drug outcomes. A study of young, homeless adolescents found that reasons for change were associated with later change, and desire and ability statements against change were associated with later maintenance of substance-use behaviors (Baer et al., 2008). Similarly, Walker, Stephens, Rowland, and Roffman (2011) found that desire and reason statements were associated with positive marijuana treatment outcomes for adults up to almost three years. Along with frequency and subcategory, the strength of change talk also has been predictive of positive client outcomes for alcohol users (Amrhein et al., 2003; Campbell, Adamson, & Carter, 2010) and cannabis users (Strang & McCambridge, 2004), which suggests that frequency alone paints an incomplete picture. Therefore, it is possible to predict behavioral outcomes from the frequency, subcategory, and strength of client change talk within individual MI encounters.

Clinician behaviors predict change talk. Although MI theory predicts that ambivalent clients naturally will express both change talk and sustain talk (Miller & Rollnick, 2002), the relative proportions of those behaviors appear to be influenced by the specific in-session behaviors of clinicians, and not only by pre-existing client factors. For example, MI-consistent clinician behaviors have been associated with increased client change talk about substance-use target behaviors (Catley et al., 2006; Gaume et al., 2008; Gaume et al., 2010; Moyers et al., 2007; Moyers et al., 2009; Vader et al., 2010). Conversely, MI-inconsistent behaviors have been associated with increased sustain talk (e.g., Gaume et al., 2010; Moyers et al., 2009).

Moreover, there is some limited evidence that clinicians can actively manipulate client behaviors in session, including change talk. An MI-based study of college drinkers showed that clinicians could influence the percentage of change talk obtained from clients within a psychotherapeutic session, simply by attending selectively to it (Glynn & Moyers, 2010). This finding suggests that clinicians—by maintaining an MI-consistent style and reinforcing only language in favor of change—might have the power to augment change talk and influence client outcomes.

***Conclusions*.** A small body of research evidence now supports all three paths of the hypothesized causal chain for change talk. Partial (Vader et al., 2010) and full (Moyers et al., 2009) support for mediation now has been found within two studies, which is an important preliminary step toward showing that change talk is a causal mechanism of individually delivered MI (e.g., Kazdin, 2007).

Client change talk in groups. Although change talk has been studied in an individual therapeutic setting, a literature review yielded just one study that explored

client commitment language—a type of change talk—in groups. Those authors (Engle et al., 2010) found that the strength of group-level commitment language to change marijuana use, as rated by a commitment-language coding system, predicted actual changes in individual client marijuana use at post-session and 12-month follow ups. Although this finding is consistent with the individual MI literature and suggests that group-level change talk might function similarly in group MI as in individual MI, the relations between in-session behaviors, client change talk, and outcomes in groups remain nearly unstudied.

Blending Motivational Interviewing with Other Treatment Elements

Although relatively little is known about how MI interacts with other treatment elements or packages, it has been combined with them both concurrently and sequentially in clinical practice (e.g., Project COMBINE; Miller, 2004), sometimes with surprising or null results (e.g., Carroll et al., 2006; Donovan, Rosengren, Downey, Cox, & Sloan, 2001; Miller, Yahne, & Tonigan, 2003; Moyers & Houck, 2011). Further research is needed to determine if, when, and how MI can best be combined with other therapies (e.g., Carroll et al., 2006), particularly for those that conflict theoretically with MI (e.g., confrontation, education, or advice-giving). This section explores one treatment element (feedback) and one manualized MI-derived intervention (Motivational Enhancement Therapy), which have been linked to beneficial client outcomes.

Feedback. One promising common-factor therapeutic element that has been studied across diverse interventions is personalized normative feedback. Clinicians provide feedback in MI (after obtaining client permission) and in other therapies (generally without permission) as a way to build discrepancy between client behavior and

cherished values or goals, and to heighten perceptions of personal risk resulting from the target behavior (e.g., Apodaca & Longabaugh, 2009; Miller & Sanchez, 1994). Feedback typically contrasts an individual client's assessment scores, medical test results, or perceptions about a behavior with those of a similar normative sample—for example, a client might learn that she ranks in the 90th percentile of same-sex marijuana smokers, shows an elevated GGT score and early signs of liver damage due to alcohol use, or uses condoms only half as often as her same-age peers.

Alone or in combination with other interventions, personalized feedback has been shown to correlate with later client behavior change. Even simple computer-generated or mailed feedback has been associated with client change, particularly reductions in alcohol use (Dimeff & McNeely, 2000; Walters, Bennett, & Miller, 2000; Walters, Vader, Harris, Field, & Jouriles, 2009; White et al., 2006), which suggests that feedback is a robust and useful clinician tool for supporting client change. Importantly, general educational approaches typically have not been efficacious in changing behavior, and thus feedback apparently must be personalized to be effective (e.g., Larimer & Cronce, 2002).

Motivational Enhancement Therapy. Motivational Enhancement Therapy (MET) refers to a manualized, MI-based intervention that includes substantial assessment and personalized-feedback components. MET originated as the Drinker's Check Up (Miller, Sovereign, & Krege, 1988) and later was employed as one of the three treatment conditions in Project MATCH (e.g., Miller, Zweben, DiClemente, & Rychtarik, 1994), but has been adapted extensively for use in other settings. Like MI, MET is considered an evidence-based treatment for substance-use disorders, and is brief and cost-effective to

deliver (CSAT, 1999; 2006; National Registry of Evidence-based Programs and Practice, 2007).

Also like MI, the mechanisms of MET remain unknown. Because MET is a combination of MI and feedback, and both elements show evidence of efficacy in improving treatment outcomes, it is unknown which element (or both, or neither) might be driving the efficacy of MET.

Several recent studies have examined MET interventions with a goal of disentangling the effects of MI components from feedback. Findings from one meta-analysis (Lundahl et al., 2010) suggested that MET actually might be more powerful than "pure" MI, although more evidence reportedly was needed to draw that conclusion. Consistent with this idea, a recent study of heavy-drinking college students concluded that MI plus web-based feedback was more effective than MI alone or feedback alone in improving alcohol outcomes (Walters, Vader, Harris, Field, & Jouriles, 2010), and that MI-consistent clinician language was related to client change talk (Vader, Walters, Prabhu, Houck, & Field, 2010). White et al. (2006) found that MI plus written feedback added no benefits over written feedback alone in reducing substance abuse in university students, which suggested that the feedback and *not* the MI was primarily responsible for changes in substance abuse. A study of adolescent marijuana users found that a school-based MET marijuana intervention was associated with reduced marijuana use at three months, but did not differ from a control group (Walker, Roffman, Stephens, Berghuis, & Kim, 2006). Thus, although MET is promising an intervention for young people, exactly how, why, and when it works remains poorly understood.

Translating MI and MET to Groups

MI in groups has been described for over a decade as "promising," but little progress has been made in empirically evaluating these groups for efficacy and potential mechanisms. Because MI and MET originated in an individual psychotherapeutic format and only recently have been applied to group settings, it is unclear whether MI translates equally well to a group-delivered format; if so, which elements of MI might be the most important to preserve; and how MI might best be combined with elements of other treatments, given that "pure" MI seems not to have been attempted yet in groups.

Perhaps reflecting these uncertainties, findings for the efficacy of group MI and MET have been mixed, although the majority have yielded positive results across various settings, formats, and target behaviors (Feldstein Ewing, Walters, & Baer, 2012; Walters, Ogle, & Martin, 2002). Those results are discussed for both adults and young people.

Adults. Four studies have examined group MI or MET in adults. First, a pilot study of group MI in adult inpatients offered descriptive analyses of a group that combined MI-based exercises with CBT and relapse-prevention exercises (Van Horn & Bux, 2001), which concluded that the materials were acceptable to clients. Second, a six-session intervention designed to enhance motivation to engage in other treatment was tested with adult DWI offenders, and was related to a doubling of the retention rate relative to standard treatment (Lincourt, Kuettel, & Bombardier, 2002). Third, dually diagnosed adults who received group MI showed decreased substance use and increased aftercare attendance, relative to a therapist-attention control (Santa Ana, Wulfert, & Nietert, 2007). Fourth, a randomized controlled trial of a "cognitive–behavioral motivational" intervention for alcohol and drugs contrasted individual- and group-

delivered formats of the same content; although the formats yielded comparable substance-use outcomes that continued to one-year follow up, the group format required 41% less clinician time (Sobell, Sobell, & Agrawal, 2009).

Adolescents and college students. A growing body of evidence has supported the use of MI-based group interventions for youth (D'Amico et al., 2011; D'Amico et al., 2012; Feldstein Ewing, Walters, & Baer, 2012), although there are several exceptions. Notably, almost all of these studies have focused upon substance use disorders as a target behavior.

One-session group interventions for adjudicated college students that incorporated an MI style with feedback or information have been correlated with decreased alcohol use and negative consequences (LaBrie, Lamb, Pedersen, & Quinlan, 2006; LaBrie, Thompson, Hutching, Lac, & Buckley, 2007; LaChance, Feldstein Ewing, Bryan, & Hutchison, 2009). However, other studies of group MI for college drinkers that included feedback have shown null results (Walters, 2000; Walters, Bennett, & Miller, 2000). A lengthier intervention for college drinkers echoed the findings of the first set of studies, and showed that participants reduced their numbers of drinking episodes and drinks per occasion relative to controls (Michael, Curtin, Kirkley, Jones, & Harris, 2006). In adolescents, a group MET (gMET) intervention for adolescent marijuana users showed a relation between the intervention and improved marijuana outcomes at two follow-up points (Engle, Macgowan, Wagner, & Amrhein, 2010). It is unclear why seemingly similar interventions in similar populations would yield such diverse results across studies, although one possibility is that the active mechanisms have been applied inconsistently.

Other studies have examined gMET in combination with other interventions. Individual feedback combined with gMET delivered in a Greek-house setting was associated with decreased alcohol use relative to control (Larimer et al., 2001); however, the effects of individual feedback and group treatment could not be separated. Similarly, brief MI combined with a group cognitive-behavioral intervention was related to improved alcohol outcome for teens (Bailey, Baker, Webster, & Lewin, 2004), but again, the individual contributions of the two therapies could not be evaluated. Although it is interesting and important to learn whether MI may be added to other treatments to make it more potent, it does not clarify which components were effective in producing client behavior change.

Positive results of group MI and gMET also have been found in adolescents, as opposed to adults and college students. For example, gMET was associated with improved drinking outcomes at three- and six-month follow-ups, relative to educational and brief-information controls (LaChance, Feldstein Ewing, Bryan, & Hutchison, 2009). Likewise, an intra-family MET intervention for adolescents and their relatives was associated with reductions in risky substance use and other behaviors (Slavet et al., 2005).

Individual MI and MET for safer sex. Several studies have addressed safer-sex behaviors with adolescents in an individual setting, with a primary focus upon people with sexually transmitted infections or HIV/AIDS.

Overall, these interventions have shown promise in improving safer-sex outcomes. Most have focused upon increasing condom use (e.g., Chen, Murphy, Naar-King, and Parsons, in press; LaBrie, Pedersen, Thompson, & Earleywine, 2008; Shrier et

al., 2001), but other behaviors have been explored as well. First, an individual intervention that targeted both safer sex and alcohol use in college women showed that MI was associated with increases in the use of effective contraception and decreases in risk for alcohol-exposed pregnancy (Ingersoll et al., 2005). Second, MET for HIV-positive young people reduced both viral load and unprotected sex acts, relative to control (Naar-King et al., 2006). Third, a brief, MI-based decisional balance was associated with increased condom use a month later in heterosexual men (LaBrie et al., 2008). Fourth, a review paper showed some positive results for various formats of MI-based safer-sex interventions, although for other formats results were contradictorily negative, and the authors were not able to explain the discrepancy (Lopez, Tolley, Grimes, & Chen-Mok, 2011). Finally, some MET interventions have been delivered successfully via computer, including for reducing births in adolescents (Barnet et al., 2009), improving HIV-prevention knowledge, and carrying condoms (Kiene & Barta, 2006). On the whole, individual interventions for safer sex seem to be efficacious, and perhaps they can help guide research on group-delivered MI and MET for this target behavior.

Group MET for safer sex. Very few studies have addressed the efficacy of gMET for safer-sex behaviors, but two have focused upon adults and two upon detained (i.e., temporarily incarcerated) adolescents. A small-group MI session for heterosexual African American men was associated with greater improvement in HIV-prevention behaviors, relative to a non-MI group (Kalichman, Cherry, & Sperling, 1999). Similarly, an "informational–motivational" (i.e., behavioral skills training combined with MET) group for HIV prevention in high-risk African American women was related to gains in risk-prevention knowledge and strengthened HIV-prevention intentions, compared to a

health-promotion control (Carey et al., 2000). In another paper from that project, it was reported that the intervention changed both risk perception and HIV-prevention behaviors at three months, including condom use, communication with partners, reduced substance use before sex, and several attitudinal or intentional changes on measures (Carey et al., 1997).

In detained adolescents, a study of gMET for safer sex showed that MET was related to reduced condom non-use after release, compared to control, although only for participants who were not depressed (Rosengard et al., 2007). Two publications about the same large-scale project compared the efficacy of three conditions of a single-session safer-sex gMET intervention for detained adolescents (Bryan, Schmiege, & Broaddus, 2009; Schmiege, Broaddus, Levin, & Bryan, 2009). They found that "risk-reduction" conditions based upon the Theory of Planned Behavior (including condom norms, attitudes, and intentions), with or without MET, were associated with better safer-sex outcomes at three months than an information-only condition, but that they did not differ from one another in terms of outcomes. However, the combined intervention was more effective than the other two conditions at changing hypothesized mediators—that is, attitudes, intentions, self-efficacy, and perceived norms (Schmiege et al.).

The factors that might be responsible for positive outcomes from gMET for safer sex, as for other MI-related interventions, are not well understood. However, despite the evidence for gMET being extremely limited, the findings are consistent, and so gMET shows promise as an intervention to reduce sex-related risk in adults and adolescents. Moreover, given the consistent findings about changes in self-report measures of safer-sex attitudes, intentions, self-efficacy, and norms following gMET interventions, it seems

particularly important to relate these constructs to more readily observable phenomena. Because client change talk has been described as a signal of client intention to change (e.g., Apodaca & Longabaugh, 2009), it offers an appropriate starting place for doing so.

Research Gaps

Despite many unknowns, MI-based interventions appear to be helpful to clients in changing problematic health behaviors, both individually and in groups. Preliminary ideas about why and how these interventions work have been hypothesized, and research has begun on identifying and testing the possible causal mechanisms of individual MI and its variants. Individual MI appears to stand on solid ground for many target behaviors.

However, much less is known about group-based MI, including which therapeutic elements it should include. That approach generally provides a nod to MI principles and is likely to be delivered in combination with other, more established group treatments. However, when discrepant treatments are combined, even those that are evidence-based, the risk always exists that they will be ineffective or confusing to clients and clinicians (e.g., Moyers & Houck, 2011). Nevertheless, the need for cost-effective and evidence-derived (if not evidence-based) treatment in the clinical setting means that group MI is being delivered to clients now, despite a limited understanding of whether, when, why, and how it might operate. Clients require better treatments that have been developed and tested using the best methods, and only researchers can begin to fill these gaps. As the National Institute on Drug Abuse (2003) has recommended, future research should "study the mechanisms of action of group therapy" in order to understand how groups operate and what the implications are for treatment.

Thus, the logical next steps are to evaluate not only the efficacy and effectiveness of group-delivered MI-based treatment, but also to study potential relational and technical factors that might serve as "active ingredients" of these treatments. To examine the individual paths of the mediational model that has been created for individually delivered MI is a starting point, and it could be accomplished by first examining clinician speech for hypothesized causal elements, and then relating those to client language during treatment sessions as well as outcomes.

Purpose, Specific Aims, and Hypotheses

The purpose of this study was to evaluate the during-group speech of clients and its relation to client outcomes, based upon a gMET intervention that focused upon safer-sex behaviors of detained adolescents. Specific aims were to (1) describe clinician and client speech using an MI coding system, and (2) predict distal safer-sex outcomes from group-level client change talk (and other relevant variables, including client sex, clinician effect, and baseline safer-sex score).

The following primary hypotheses were tested:

- Hypothesis 1a. Greater levels of group Percentage Change Talk will predict better Risky Sex Index scores for individuals at three months.
- Hypothesis 1b. Greater levels of group Percentage Change Talk will predict better Risky Sex Index scores for averaged groups at three months.
- Hypothesis 2a. Greater levels of group Percentage Change Talk will predict better Safer Sex scores for individuals at three months.
- Hypothesis 2b. Greater levels of group Percentage Change Talk will predict better Safer Sex scores for averaged groups at three months.

Method

Data for this study were derived from Dr. Angela Bryan's ongoing parent study, Project MARS (Motivating Adolescents to Reduce Sexual Risk; NIAAA #R01AA013844-05A1). The primary aim of Project MARS was to evaluate the relative efficacy of three versions of a group Motivational Enhancement Therapy (gMET) intervention in improving proximal and distal risk-reduction outcomes for three target behaviors (safer sex, alcohol, and marijuana). That study was a five-year project with projected total sample of 520 detained adolescents.

Participants

Group leaders (clinicians). Intervention leaders were seven adult clinicians (three women and four men) who had completed at least a baccalaureate degree. All clinicians received reading material and training in the basics of MI and the manualized intervention, observed experienced group leaders, conducted mock interviews that were reviewed for proficiency, and met a minimum quality threshold before conducting any groups (Project MARS, 2010). Leaders also received weekly clinical supervision throughout the study from Dr. Sarah Feldstein Ewing, a licensed clinical psychologist and member of the Motivational Interviewing Network of Trainers; supervision techniques included recording review and discussion of difficult areas based upon leader self-report and intervention-fidelity ratings.

Recruitment. Client participants were recruited from the Bernalillo County Juvenile Detention Center in Albuquerque, NM, USA. Adolescents provided informed assent, and their parents provided informed consent prior to scheduling or initiation of study procedures. Inclusion criteria were current detention with an anticipated duration of

less than one month, being able to speak English, having parental consent to participate, and agreeing to be contacted at four follow-up points (three, six, nine, and twelve months post-intervention).

Adolescents (clients). Participants from which the sub-sample was drawn were 200 adolescent members of 58 group interventions of one to six participants, who ranged in age from 14 to 18 years (M = 16.02, SD = 1.06). More participants were male (76.5%) than were female (23.5%). Most participants (77.5%) identified as sexually attracted to people of the opposite sex, 5.5% to the same sex, and 12.5% to both sexes, and 4.5% elected not to answer the question due to discomfort. Nearly all (94%) reported ever having had vaginal or anal intercourse; mean age of first sex was 12.9 years (SD = 1.8 years), and the mean number of lifetime sexual partners was 11.81 (SD = 15.4). In the three months between Baseline and Three Month, many participants reported having used substances prior to sexual intercourse at least "sometimes" (35% for alcohol and 44% for marijuana). Racioethnically, the adolescents represented in the study were very diverse, and identified as Hispanic (66%), Caucasian (17%), American Indian/Native American (5.5%), African American (5.5%), Mixed (5.5%), and unknown/unspecified (0.5%). Two thirds of participants reported that they qualified for either a free (51.1%) or reduced-price (16.5%) school lunch, which served as a proxy for low socioeconomic status. Although reasons for detention varied, the most common (37.5%) was assault or fighting.

In the sample retained for analysis (n = 45 individuals), participants were similar to the full sample on many demographic variables, including racioethnic background, age, and income level. However, they differed in several ways potentially related to outcome variables: Fewer participants (17.80%) were female, more participants (86.7%)

identified as sexually attracted to people of the opposite sex, fewer (93.3%) reported ever having had vaginal or anal intercourse, average age of first sex was higher (13.14 years, *SD* = 1.61 years), and the mean number of lifetime sexual partners was lower (*M* = 9.89, *SD* = 15.06). Alcohol use prior to sexual intercourse was similar to the original sample (35%), but marijuana use before sex was less prevalent (29%).

Protection of human subjects. The parent study was overseen by the University of New Mexico (UNM) Main-Campus Institutional Review Board (IRB). This author and the two coders held IRB approval as Project MARS investigators, and thus were approved to observe interventions and access study data.

The present study, which was a secondary analysis of existing audio and numeric data, included the risk of possible breach of confidentiality through review and coding of group recordings. No further direct benefits to participants were anticipated above those offered by the parent study. Risks that resulted from coding audio recordings of clients' in-session speech were not expected to pose significantly increased risks beyond those inherent in Project MARS, and no adverse events resulted from the addition of these procedures. Furthermore, using existing data efficiently through secondary analysis allowed answering of additional questions while avoiding additional risk to participants.

Data storage. Audio recordings for Project MARS were stored on a secure, limited-permissions folder on the Center on Alcoholism, Substance Abuse, and Addictions (CASAA) network. Other electronic data were stored within a different folder on this network, and numeric links (not other identifiers) connected participant data to group audio recordings. Only trained study staff held access to those electronic files.

Coding led to the generation of two additional types of de-identified data: Electronic text files (i.e., .parse, .casaa, and .globals formats) and paper-copy coding sheets. Text files were stored in a separate folder CASAA network, apart from the audio recordings, and coding sheets were stored in a locked file cabinet in a locked office at CASAA prior to being entered electronically for analysis. In accordance with American Psychological Association ethical guidelines, files will be retained for at least five years following publication. Then they will be destroyed by deletion (electronic) or cross-cut shredding (paper).

Materials

Measures. In the parent study, assessment batteries were be administered to participants via laptop computer at Baseline and at all follow-up points (i.e., PostTest, Three Months, Six Months, Nine Months, and Twelve Months). The specific measures included in the Project MARS battery varied by time point, but the measures included in the present study were identical and administered at Baseline (i.e., immediately prior to the group intervention) and at Three Months post-intervention.

Sexual History. The Sexual History questionnaire is a collection of items derived from Dr. Bryan's previous studies on safer sex in adolescent populations (Project MARS, n.d.), and does not represent a formal scale. Two items about the frequency of sex and use of condoms were multiplied to form the Risky Sex Index. This item showed construct validity through its association with theoretically relevant social-cognitive mediators (e.g., self-efficacy about condom use), and was responsive to intervention in a randomized trial (Schmiege et al., 2009). Two other items about safer-sex behaviors during the previous three months were averaged to create the Safer Sex variable, which

indicates the degree to which the participant reported in using condoms or other birth control during the past three months. (See Key Variables, below, for item details.)

Key Variables

The following key variables served as grouping variables, predictors, or outcome variables. (See Table 1 for a summary.)

Past-three-months sexual intercourse (Intercourse). The Intercourse variable indicated whether the participant reported having sex during the past three months at the Three Month follow-up. This item is (Project MARS, n.d., p. 9):

- Have you had sexual intercourse in the *past three months?*

The intercourse item was dichotomized into 0 = "no" or 1 = "yes" responses and used as a grouping variable: Primary study analyses for individuals were performed only for participants who responded affirmatively to the Intercourse item at Three Months.

Group. The Group variable represented each therapy group by only one case for group-level analyses. Cases were designated as either 0 = "not selected" or 1 = "selected" and used as a grouping variable. No values for outcomes variables were missing from selected cases, unless every member of a group had a missing score for that item.

Percentage Change Talk. Percentage Change Talk is a summary variable that is calculated as the frequency of change talk divided by the sum of the frequency of change talk plus the frequency of sustain talk; that is, % CT = (CT / [CT+ ST]). Possible values can be expressed as either decimals (ranging from 0 to 1), or as percentages (ranging from 0% to 100%). That variable was calculated over the entire gMET session. Percentage Change Talk was used as a predictor variable for both safer-sex outcome

variables (i.e., Risky Sex Index and Safer Sex). Normative data for group-level change talk were not available from the literature.

Self-reported risky sex (Risky Sex Index). The Risky Sex Index had been used previously in Dr. Bryan's studies (e.g., Schmiege et al., 2009) to estimate the degree of sexual risk of participants. The Risky Sex Index is calculated by multiplying the average frequency of sexual intercourse (1 = "A few times a year" to 6 = "Almost every day") by the reverse-scored frequency of condom use (5 = "Never" to 1 = "Always"); scores can range from 1 to 30, with higher scores representing greater sexual risk. Items included (Project MARS, n.d., p. 19):

- On average, how often do you have sexual intercourse?
- How much of the time have you used condoms when you've had sexual intercourse?

Those variables were calculated in two ways: For individuals based upon their individual scores (i.e., Risky Sex Index–Three Month–Individual; RSI-3M-I), and for mean scores of participants who shared a gMET intervention at Baseline (i.e., Risky Sex Index–Three Month–Group; RSI-3M-G).

Self-reported sexual behavior (Safer Sex–Baseline, Safer Sex–Three Month). The Safer Sex–Baseline and Safer Sex–Three Month variables represented safer-sex behaviors during the preceding three months, which were derived from an average of two items from the Sexual History questionnaire at the respective time point (Project MARS, n.d., p. 22):

- In the past three months, how much of the time did you use condoms when you had sexual intercourse?

- In the past three months, how much of the time did you use some other kind of birth control when you had sexual intercourse?

They were chosen to emphasize behaviors consistent with prevention of HIV and pregnancy, which are prevalent concerns in this population (e.g., Tolou-Shams, Stewart, Fasciano, & Brown, 1999; Whaley, 1999). For individual analyses, those variables were calculated only for participants who had at least one sexual partner between Baseline and Three Month time points (i.e., who responded "yes" on the Intercourse variable). Possible averages ranged from 1 = "Never" to 5 = "Always," with higher scores representing greater employment of safer-sex behavior. As with the Risky Sex Index, those variables were calculated for individuals (i.e., Safer Sex–Baseline–Individual) and for averaged groups (i.e., Safer Sex–Baseline–Group).

Client Sex. Some evidence has suggested that male and female adolescents differ in their sex behaviors (e.g., age of first intercourse, number of partners, HIV knowledge, safer-sex intentions, sexual impulsivity; Canterbury, Clavet, McGarvey, & Koopman, 1998), with female adolescents generally showing greater knowledge and lower risk. Similarly, boys and girls respond differently to gMET interventions (S. Feldstein Ewing, personal communication, October 27, 2010). Thus, client sex was used as a predictor of safer-sex outcomes. Client sex was matched with clinician sex in all interventions.

Clinician. Limited past research on change talk has shown that clinicians who are trained together and who achieve proficiency standards prior to leading interventions still may obtain significantly different amounts of individual change talk from their clients (i.e., 49%–73%, collapsed across clients; Glynn & Moyers, 2010, p. 68). The seven clinicians in the study were included in regression equations (via unweighted effect

codes) to evaluate differences in client outcomes based upon which clinician led the group.

Coding instrument and software. Behavioral coding of the audio-recorded gMET interventions (see Procedure, below) was accomplished using the Motivational Interviewing Skill Code (MISC 2.5; Houck et al., 2010) in combination with the CASAA Application for Coding Treatment Interactions software (CACTI; e.g., Glynn, Hallgren, Houck, & Moyers, 2012).

Motivational Interviewing Skill Code (MISC 2.5). The MISC 2.5 (Houck et al., 2010) is a coding system developed at UNM for the sequential analysis of psychotherapeutic interactions. It combines aspects of two previous coding systems for motivational interviewing: The Motivational Interviewing Skill Code (MISC 2.1; Miller, Moyers, Ernst, & Amrhein, 2008) and the Sequential Code for Observing Process Exchanges (SCOPE; Martin, Moyers, Houck, Christopher, & Miller, 2005). MISC 2.5 was used to reliably code hundreds of therapeutic interactions for Dr. Theresa Moyers' five-year trial of training clinicians in motivational interviewing (Project ELICIT; NIDA #1R01DA021227), with inter-rater reliabilities upward of "good" to "excellent" (i.e., intraclass correlations [ICCs] above 0.6; Cicchetti, 1994) for 13 of the 14 summary scores (Glynn et al., 2012).

The MISC 2.5 allows for the parsing (dividing) and sequential coding of client and clinician utterances, as well as for coding client and clinician global ratings. MISC 2.5 can be used with either session transcripts and recordings, or with specialized coding software (see CACTI, below). The sequential nature of the system preserves the order of utterances within the session, which allows for evaluation of the therapeutic process.

Although sequential hypotheses were not examined in the present study, they could be tested in future, exploratory analyses using existing data.

Coding with MISC 2.5 requires two passes per recording. In the first pass, coders parse client and clinician speech into individual utterances (i.e., "thought units;" Houck et al., 2010, p. 1). During this pass, coders also use paper coding sheets to take notes about the session and then to assign one global rating of the client (Self-Exploration) and six global ratings of the clinician (Acceptance, Empathy, Direction, Collaboration, Autonomy Support, and Evocation). Global ratings fall on a five-point Likert-type scale, with "1" as the default for the client rating and "3" as a default for the clinician ratings.

In the second pass, coders assign client or clinician behavior codes to the pre-parsed utterances. Generally, coders avoid coding recordings that they parsed previously, although doing so is unavoidable with double-coded recordings for reliability. Client behavior codes include one follow-neutral category, seven sub-categories of change talk, and seven sub-categories of sustain talk. Clinician behavior codes include three neutral behaviors, four question types, eight reflection types, six MI-consistent behaviors, and five MI-inconsistent behaviors. Any utterance may be assigned a "no code" rating if it is inaudible or irrelevant to the session (e.g., speech from third parties, breaks or non-session-related behaviors, long periods of silence). When utterances are ambiguous, the MISC 2.5 manual (Houck et al., 2010) specifies "decision rules" that help guide coders resolve the coding of ambiguous utterances and improve reliability.

In addition to examining client and clinician language and testing sequential hypotheses, MISC 2.5 allows for calculating estimates of clinician fidelity through the use of clinician summary variables. Those can be obtained through various MI coding

systems, and have been used in previous studies (including large clinical trials) to evaluate clinician fidelity to MI (e.g., Gaume et al., 2008; Gaume et al., 2010; Glynn et al., 2012; McCambridge et al., 2011; Moyers & Martin, 2006; Moyers et al., 2009). Relative to individual frequencies, summary variables tend to have the disadvantage of being less precise, but the advantage of showing higher inter-rater reliability.

The CASAA Application for Coding Treatment Interactions, version 9.7 (CACTI). The CACTI software (e.g., Glynn et al., 2012) is a free, open-source program that was developed for concurrent use with the MISC 2.5 coding system. It is a two-pass "on-the-fly" parsing and coding program for digital (.wav) audio recordings that was designed to replace transcription, minimize data entry, and increase the efficiency of MISC 2.5.

CACTI users click buttons on the graphic user interface to signal the beginnings and endings of parses, to assign codes to pre-parsed client and clinician utterances, or to submit global ratings of clients and clinicians. The three modes of the program create three file types (parsing files = .parse, sequential-coding files = .casaa, and global ratings = .globals), which can be opened as plain-text (.txt) files and imported into statistical software for analysis.

When used with MISC 2.5 by experienced coders for individual psychotherapy sessions, parsing and coding global ratings using the CACTI software session has taken approximately 1.25–1.5 times the length of time of the therapy session, and sequential coding approximately 1.0–1.25 times the length of the session (e.g., a one-hour recording would take 60–90 minutes to parse, and 60–75 minutes to code sequentially). Recordings that are of poor audio quality, have frequent overlap between clinician and client speech,

or include rapid or unfamiliarly accented speaking styles can increase these time estimates.

In this study, because of the complexity and rapidity of the group interactions and the extended length of the session (i.e., about two hours; *M* = 1:57:18, *SD* = 15:05), coders reported that group sessions took much longer to code, led to greater fatigue, and necessitated more frequent breaks relative to individual sessions. Indeed, utterances per session averaged 966.88 (*SD* = 204.44) across sequentially coded sessions in this study, which greatly exceeded the overall count from the "pure" sample from Project ELICIT (i.e., *M* = 228.49, *SD* = 91.67). However, the rates of utterances were more similar, with *M* = 7.07 per minute (*SD* = 4.46) in this study, and *M* = 6.02 (*SD* = 1.92) in Project ELICIT (S. Lloyd Rice and J. Houck, personal communication, March 20, 2013).

Procedure

The Project MARS parent study included one initial intervention and data-collection session and follow-up data-collection sessions at three, six, and nine months post-intervention. Because intervention effects were expected to be strongest closer to the session, only data from the Baseline and Three Month visits were used to test the hypotheses of the present study.

Baseline visit. During the initial visit, clients (1) completed the Baseline assessment battery individually on laptop computers, (2) provided urine and saliva samples, (3) participated in a group Motivational Enhancement Therapy (gMET) intervention, and (4) completed PostTest measures individually on laptop computers. Only measures from the Baseline assessment battery (i.e., Demographics and Sexual

History questionnaire), and audio recordings from the gMET intervention were used from this visit.

Intervention structure and randomization. As part of the initial visit, all clients participated in a single gMET intervention that included single-sex cohorts of two to six adolescent clients and one same-sex adult clinician. Groups were randomly assigned via a random-number table to one of three conditions (described in the next paragraph). Participants were assigned to groups depending upon their sex and the date fitting their stay at the Detention Center. Because more boys than girls were detained at the Detention Center, female groups were smaller and were conducted less frequently.

Groups were assigned to one of three conditions, which focused upon different target behaviors: Sex Only, Sex + Alcohol, and Sex + Alcohol + Marijuana. The content and duration of the group intervention varied, but all conditions were manualized and designed to average 105 minutes. Group leaders were instructed to attend only to change talk related to the target behavior(s) for the appropriate condition.

Overview of the gMET intervention. The gMET intervention was a group-delivered therapeutic session with a primary focus of reducing sexual risk. Consistent with an MET framework, the intervention was designed to present specific content elements (i.e., safer-sex education and personalized normative feedback) within an MI-adherent clinician style (Project MARS, 2010).

All groups incorporated the following activities (with approximate times in parentheses:

- Introductions and generation of group rules (10 minutes)
- Definitions exercise (10 minutes)

- AIDS flash-card exercise (10 minutes)
- Provision of peer norms, and readiness rulers (15 minutes)
- Self-affirmation exercise (10 minutes)
- High-risk situations movie, plus discussion (20 minutes)
- Hands-on condom and dental-dam demonstrations, discussion of how to obtain condoms, and safer-sex sequencing exercise (25 minutes)
- Summary and wrap-up (5 minutes)

Three Month follow-up visit. At three months post-intervention, Project MARS participants were re-contacted and asked to schedule an individual appointment at one of three locations: an annex at the Detention Center, the participant's home, or the Project MARS lab at UNM. An assessment battery was administered to each adolescent individually via laptop computer. Only data from the Sexual History questionnaire were used from the Three Month assessment battery. The three-month follow-up rate of 88% in this sample was favorable, and was even higher than rates from Dr. Bryan's previous studies with adolescents in this setting (e.g., 60% in Schmiege et al., 2009; 83% in SHARP [T. Callahan, personal communication, April 15, 2011]).

Coding of audio recordings. Preparation for coding with the MISC 2.5 coding system and the CACTI software (described above in Materials) was a multi-step process.

***Audio preparation*.** Digital recordings of groups of clinicians and clients from the gMET intervention were prepared for coding by the investigator. All digital audio recordings from Project MARS required conversion from .wma to .wav format for compatibility with the CACTI software, which was accomplished using a trial version of GoldWave Audio Editor software (GoldWave, Inc., 2013). Also, because 12 sessions

were recorded in two segments due to restroom breaks, those files required recombination prior to conversion. Two recordings had fully inaudible portions, but overall, audio was judged to be of acceptable quality for coding.

Specification of target behavior. Because the three conditions of Project MARS included between one and three target behaviors, only safer-sex precautions—targeted in all three conditions—was chosen as the focus for this study. This target behavior included discussion of reductions in substance abuse implemented specifically in support of safer-sex behaviors (e.g., "I want to smoke less weed so I remember to use a condom," or "I'm going to cut down to three beers so I don't accidentally get pregnant"). For instances in which change talk clearly was occurring but it was ambiguous whether the target behavior was safer sex or alcohol/marijuana use, a decision rule was created and added to the coding manual; the goals of that rule were to aid coders in rating the utterance reliably, and to err on the side of capturing safer-sex change talk instead of excluding it. Discussion of changing the other intervention target behaviors (e.g., "I should stop drinking so I don't hurt my liver" or "Marijuana makes it harder to drive") therefore was not coded as change talk, and instead fell within the Follow/Neutral category. Percentage Change Talk was not found to vary among the three conditions ($F(2, 31) = 0.017$, $p = 0.983$).

***Coders and coder training*.** Coders were two paid graduate-student assistants who had worked on a previous project within the coding lab. Although coder training typically involves one semester of didactics, readings, written exercises, and individual and group coding practice, training time was greatly reduced for this project because the two coders were already reliable in the MISC 2.5 for rating individual psychotherapy sessions with

adult substance-using clients. Thus, retraining coders involved applying the system to group interactions, a safer-sex target behavior, and an adolescent population, and then defining additional decision rules to help coders reach consensus and aid in achieving acceptable reliability.

The coding team met approximately every other week during study coding to resolve concerns and practice coding together. Formal reliability checks were scheduled to be performed periodically throughout the study to minimize coder drift and allow for mid-study correction. Unfortunately, this was not feasible because the minimum number of sequentially coded recordings required to evaluate inter-rater reliability meaningfully using ICCs (i.e., approximately five) was not available until the end of the study. (See Inter-Rater Reliability below for details.)

Exclusion of sessions and randomization. The coding universe of sessions from Project MARS included 58 audio-recordings of gMET interventions that were available at the time of coder randomization. One file was corrupted and therefore was discarded; of the 57 remaining sessions, 9 were randomized via a random-number generator to be used for reliability training or group practice, and the remaining 48 were assigned masks for parsing and coding. Sessions were randomized to be divided approximately equally between the two coders for parsing, and then coders were intended to sequentially code the opposite sessions, with 20% overlap planned to assess reliability. After excluding participants (see Exclusion of Participants, below), 27 intervention groups were represented in the final analyses for individual outcomes.

Because of investigator underestimation of the additional burden of coding group sessions, difficulties with fund disbursement, and fluctuating coder availability for the

project, coders were unable to parse or code as many sessions per week as anticipated. The two coders also had unequal availability for coding and progressed at discrepant rates, which was inefficient because it necessitated one coder waiting for the other to complete parsing in order to begin coding. To allow coding to progress, the investigator instead parsed all recordings designated for coding (n = 44, after exclusion of an inaudible session). Coders then began double-coding a reliability sample, and later completed as much of their original coding assignment as possible. Because available coding time was diminished, 17 parsed recordings did not undergo sequential coding, and 3 training recordings were included in the reliability sample despite initial plans to select only from the designated non-training sessions, which might have led to underestimation of overall reliability. (See Figure 1 for a flowchart of audio-recording selection and exclusion.)

Results

Preliminary Analyses

Data preparation. As planned, preliminary analyses were performed to ensure data quality and adherence to accepted statistical standards. Those included plotting data for descriptive purposes, testing for violations of the assumptions of multiple regression, and evaluating multicollinearity and internal consistency.

Visual inspections were made of the P-P plots of standardized residuals, as well as the scatterplots of continuous predictors against each dependent variable, predicted values against residuals, and standardized residuals against continuous predictors (with Loess lines), and ID number against standardized residuals. Those inspections provided some evidence of linearity, normality, homogeneity of variance, and independence in all four models, although distributions were sparse and slightly aberrant, likely due to small sample size.

Tests of multicollinearity were performed for all four models using the Variance Inflation Factor (VIF), which estimates the severity of the overlap between an independent variable of interest and the other independent variables in a model (e.g., O'Brien, 2007). VIFs ranged from 1.0 to 1.2 (i.e., did not exceed recommended cutoffs of 10), which suggested non-redundancy in predictors and acceptable numbers of predictors per final model.

Internal consistency was calculated for the Risky Sex Index–Three Month–Individual and Safer Sex–Three Month–Individual variables using both simple correlations and Cronbach's alpha (Cronbach, 1951). The Risky Sex Index was unexpectedly negative, at $r = -0.323$. The Safer Sex variable yielded $r = 0.523$, which

approached acceptable strength without being too redundant. However, those estimates should be interpreted with caution because each index included only two component items (i.e., a single correlation).

Unweighted effect coding. In the four primary analyses, the Clinician variable was coded because it was categorical and could not be evaluated directly using multiple regression. Unweighted effect coding was intended to allow testing of clinician differences, relative to the grand mean. The base group for Clinician (coded as "-1") was the clinician whose clients showed % CT closest to the grand mean; all other clinicians were coded as "1" for the code variable and "0" for the other variables.

However, the Clinician variable ultimately was excluded from final analyses because it was found to be perfectly multicollinear with Client Sex. Although high VIF does not automatically require elimination of a variable (O'Brien, 2007), the complete overlap in these models prevented the models from being statistically possible or executable in the statistical software. Only one predictor could be chosen, and Client Sex was deemed to be more important for three reasons: First, the omnibus test of Percentage Change talk failed to yield significant between-clinician differences ($F(6, 27) = 0.737$, $p = 0.625$); second, the possible effects of Client Sex were determined to be more relevant to the research question than were those for Clinician; and third, removing Clinician would preserve greater degrees of freedom.

Power. The required sample size for the most complex multiple-regression analysis planned was calculated *a priori* for power ≥ 80% using G*POWER software (Erdfelder, Faul, & Buchner, 1996; Faul, Erdfelder, Lang, & Buchner, 2007). Because effect sizes for these analyses were unavailable from previous literature, standard values

(f^2 values of Small = .02, Medium = .15, and Large = .35; Cohen, 1992, p. 157) were used to determine the range of power expected for this study. The sample size of individual clients expected to achieve adequate power ranged from 850 for a small effect, to 123 for a medium effect, to 59 for a large effect.

Exclusion of participants. The initial sample from Project MARS included 200 client participants. Participants then were excluded from statistical analyses due to one or more of the following reasons: an audio recording associated with the session was not available (n = 3), the adolescent participant was the only participant in the intervention, and therefore it did not constitute a group (n = 3), the participant did not complete Three Month questionnaires (n = 24), the participant did not report sexual intercourse between Baseline and Three Month (n = 89), or the session had not been sequentially coded (n = 36). After all exclusions, the sample comprised 45 participants (23% of initial N) that represented 27 groups (47% of initial N). For group-level outcome analyses, Risky Sex Index–Three Month–Group and Safer Sex–Three Month–Group were calculated by collapsing across all valid cases to form one unique-group score, and groups without mean outcome data for a particular variable were excluded from those analyses. Of the 34 sequentially coded groups with outcome-variable means, 32 (94% of initial N) were retained for Risky Sex Index–Three Month–Group and 27 (79% of initial N) for Safer Sex–Three Month–Group. (See Figure 2 for a participant flowchart.)

Primary Analyses

Multiple regression was selected for the primary statistical analyses employed in this study because it can determine how well the predictors estimate outcomes together, describe the relative and incremental contributions of individual variables, flexibly

incorporate many types of variables, and be analyzed and interpreted simply, relative to hierarchical analyses (e.g., Cohen, Cohen, West, & Aiken, 2003). The downside of that approach was that data were necessarily nested in this study, and violations of the assumption of independence could not be adequately assessed or addressed.

The general equation for the intended tests of all models was $\hat{y} = b_0 + b_1(Client\ Sex) + b_2\,(Clinician\ Contrast_1) + b_3(Clinician\ Contrast_2) + b_4(Clinician\ Contrast_3) + b_5(Clinician\ Contrast_4) + b_6(Clinician\ Contrast_5) + b_7(Clinician\ Contrast_6) + b_8(Percentage\ Change\ Talk) + b_9(Baseline\ Score)$. However, because of the problem of perfect multicollinearity between Clinician and Client Sex (described above in Unweighted effect coding.), Clinician was excluded from the final model. Thus, the equation tested for all models was: $\hat{y} = b_0 + b_1(Client\ Sex) + b_2(Percentage\ Change\ Talk) + b_3(Baseline\ Score)$. (Specific variables for the outcome and Baseline score for each model are outlined below.) All predictors were entered simultaneously—that is, as coequal predictors of outcomes. However, Risky Sex Index–Baseline and Safer Sex–Baseline scores were interpreted as covariates, because pretest scores had been expected to correlate highly with posttest scores.

The null hypothesis for the omnibus test of each model was that $\rho^2 = 0$, and the alternate hypothesis was that $\rho^2 \neq 0$. For tests of the effects of the individual predictors, the null hypothesis was that $\beta = 0$, and the alternate hypothesis was that $\beta \neq 0$. All analyses were conducted at $\alpha = .05$, two-tailed, and corrections for multiple comparisons were not applied because all analyses were planned *a priori*.

Hypothesis 1a. Risky Sex Index–Three Month–Individual scores were predicted from Percentage Change Talk, Client Sex, and Risky Sex Index–Baseline–Individual (as

a covariate). In that model, $\hat{y}$ = Risky Sex Index–Three Month–Individual and x_3 = Risky Sex Index–Baseline–Individual. The full model was statistically significant ($F(3, 40) = 3.090$, $p = 0.038$, $R^2 = 0.188$, Adjusted $R^2 = 0.127$), the null hypotheses was rejected, and it was concluded that the regression coefficient differed significantly from zero. Only the Baseline score significantly predicted the Three Month score ($b = 0.422$, $SE = 0.170$, $t(40) = 2.476$, $p = .018$). Results for individual predictors and correlation matrices are presented in Tables 2 and 3.

Hypothesis 1b. Risky Sex Index–Three Month–Group scores were predicted from Percentage Change Talk, Client Sex, and Risky Sex Index–Baseline–Group (as a covariate). In that model, $\hat{y}$ = Risky Sex Index–Three Month–Group and x_3 = Risky Sex Index–Baseline–Group. ($F(3, 28) = 1.009$, $p = 0.403$). The full model was not statistically significant, and the null hypothesis could not be rejected. Results for individual predictors and correlation matrices are presented in Tables 4 and 5.

Hypothesis 2a. Safer Sex–Three Month–Individual scores were predicted from Percentage Change Talk, Client Sex, and Safer Sex–Baseline–Individual (as a covariate). In that model, $\hat{y}$ = Safer Sex–Three Month–Individual and x_3 = Safer Sex–Baseline–Individual. The full model was statistically significant ($F(3, 34) = 6.421$ $p = 0.002$, $R^2 = 0.362$, Adjusted $R^2 = 0.305$), the null hypothesis was rejected, and it was concluded that the regression coefficient differed significantly from zero. Only the Baseline score significantly predicted the Three Month score ($b = 0.658$, $SE = 0.175$, $t(34) = 3.770$, $p = .001$). Results for individual predictors and correlation matrices are presented in Tables 6 and 7.

Hypothesis 2b. Safer Sex–Three Month–Group scores were predicted from Percentage Change Talk, Client Sex, and Safer Sex–Baseline–Group (as a covariate). In that model, $\hat{y}$ = Safer Sex–Three Month–Group and x_3 = Safer Sex–Baseline–Group. The full model was statistically significant ($F(3, 23) = 3.542$, $p = 0.030$, $R^2 = 0.316$, Adjusted $R^2 = 0.227$), the null hypotheses was rejected, and it was concluded that the regression coefficient differed significantly from zero. Only the Baseline score significantly predicted the Three Month score ($b = 0.915$, $SE = 0.401$, $t(23) = 2.280$, $p = .032$). Results for individual predictors and correlation matrices are presented in Tables 8 and 9.

Inter-Rater Reliability

Behavior counts and global ratings were evaluated for inter-rater reliability to ensure acceptability for use in the study.

Reliability of behavior counts. A reliability sample of approximately twenty percent of sequentially coded audio-recordings ($n = 7$; 21%) was analyzed using ICCs. Coders were trained to at least "good" inter-rater reliability (i.e., ICCs ≥ 0.6; Cicchetti, 1994) on 9 of 13 summary variables (see Table 10).

Two features of ICCs were demonstrated there: Summary variables tended to be higher in reliability than their component frequency variables, and variables that occurred relatively rarely tended to show lower reliability (e.g., for the reliability sample, mean frequency of MI Consistent = 254 vs. mean frequency of MI Inconsistent = 8). Importantly, reliability for the key variable (Percentage Change Talk) was "excellent," although its interpretation is debatable given its "fair" and "good" component variables. Closer examination of those variables showed mean frequencies for Coder 1 of Change Talk = 54.71, Sustain Talk = 19.57, and Percentage Change Talk = 73.95%, and means

for Coder 2 were Change Talk = 40.86, Sustain Talk = 15.43, and Percentage Change Talk = 73.57%. Thus, Percentage Change Talk was almost identical for the two coders within the reliability sample, but it appears to have been derived differently: Sustain Talk occurred infrequently relative to Change Talk but coders showed high agreement when it occurred, whereas Change Talk frequencies occurred more frequently and differed notably. Perhaps coders disagreed in detecting Change Talk amongst other codes (or No Code, which was prevalent in the sample), but did they also did not confuse Change Talk with Sustain Talk. Implications are explored below in Limitations.

Training recordings were not included in the original plan for calculating inter-rater reliability due to the possibility that coders would not yet be fully trained, but they were included in primary analyses to represent the sample of sessions actually used in analyses. Because reliability for the component variables to Percentage Change Talk were unexpectedly low, given "excellent" reliability for that summary variable, further exploration of the discrepancy was conducted by estimating ICCs only for non-training sessions (n = 4). In that non-training reliability sample, all but two variables achieved at least "good" reliability: Four scores improved in range (including Change Talk and Sustain Talk frequencies), one worsened (Percentage Change Talk), six remained unchanged, and two (MI-Inconsistent, Percentage MI Consistent) became uninterpretable due to negative ICCs. (See Table 10 for details.) Negative ICCs could reflect widespread disagreement (e.g., Hallgren, 2012), or just differences in detection of the low-frequency MI-Inconsistent variable. All in all, inter-rater reliability improved somewhat for this later-captured sample, which provided some support for coders having become more reliable following training.

Reliability of global ratings. The reliability of global ratings was calculated using percentage agreement. Percentage agreement is calculated by subtracting the lowest coder score from the highest, and then summing percentage absolute agreement, percentage within one integer, and percentage within two integers; ideally, the total exceeds 80%. This method is not as stringent as an ICC and has been criticized as inadequate for calculating inter-rater reliability (e.g., Hallgren, 2012), but it does allow for simple and intuitive estimation. In the present study, coders differed by no more than two ratings (on a five-point scale) on any particular reliability session, which means they achieved percentage agreement of 100% for all six global ratings within the five-session reliability sample.

Description of Client and Clinician Language, and Clinician Fidelity to MI

Client and clinician language were described using summary variables from MISC 2.5 coding for both behavior counts and global ratings. Overall clinician fidelity to MI then was assessed by comparing those summary variables to standards for "basic competence" (M = 3.5) and more advanced "proficiency" (M = 4.0) in MI suggested in the MITI 3.1.2 manual (Moyers et al., 2013, p. 29) for those variables with recommendations available.

Description of behavior counts. The sample of clinician and client in-session behaviors was described with summary variables for the 34 sequentially coded sessions. (See Table 11 for a summary.) On average, study clinicians achieved basic competence for two summary variables: Percentage MI Consistent, and Percentage Open Questions.

Description of global ratings. Coders provided global ratings for 36 unique sessions, with five (7.4%) double-coded. (See Table 12 for descriptive statistics.)

Averaged across sessions and clinicians, intervention leaders achieved "proficiency" (i.e., $M = 4.0$) on three global ratings (Acceptance, Empathy, and Direction), and approached but did not reach "basic competence" (i.e., $M = 3.5$) for all MI Spirit globals. For clients, the Self-Exploration rating was lower than the default value of 3, which could suggest that clients were low in insight about their safer-sex behaviors, or might just represent a mismatch between the coding system (designed for adults) and the adolescent participants.

Discussion

Summary of Results

Assumptions. Assumptions for multiple regression were evaluated, and there was some evidence of linearity, normality, homogeneity of variance, and independence, although small sample size probably contributed to somewhat imperfect plots. In addition, there was evidence of internal consistency for Safer Sex component items but not for Risky Sex Index, and there was no evidence of problematic multicollinearity among predictors in the final analyses.

Primary analyses. Four multiple regression analyses evaluated how well the group and individual characteristics predicted each safer-sex dependent variable at Three Months. In each analysis, the predictors were Client Sex, Percentage Change Talk, and Baseline score. Three full-model analyses were statistically significant for Risky Sex Index–Individual ($F(3, 40) = 3.090$, $p = 0.038$), Safer Sex–Individual ($F(3, 34) = 6.421$ $p = 0.002$), and Safer Sex–Group ($F(3, 23) = 3.542$, $p = 0.030$), but in each model the only statistically significant predictor of Three Month outcomes was Baseline questionnaire score. Those findings were consistent with the Baseline variables' hypothesized roles as covariates, but not supportive of alternate hypotheses regarding Client Sex and Percentage Change Talk as important predictors of outcomes.

Inter-rater reliability. Coders achieved "good" or "excellent" inter-rater reliability, as measured by ICCs, for nine of thirteen behavior-count summary variables examined; however, four ICCs in the "fair" and "poor" range were observed as well. In an exploratory analysis, reliability was improved for most variables by eliminating three training sessions from the analysis, although two variables then became uninterpretable.

Inter-rater reliability of global ratings was evaluated using Percentage Agreement, and coders were found to have agreed within two integers on all globals, for a total of 100%. Overall, the evidence for inter-rater reliability was favorable.

Intervention fidelity. The findings suggest that, on average, study clinicians approximated an intervention style and specific behaviors characteristic of clinicians who are beginning to deliver MI with fidelity. However, individual global ratings varied notably, and some sessions were rated as low as 2 (on a scale of 1 to 5), which indicated that clinicians likely did not deliver equally skillful interventions.

Discussion of Findings

The purpose of this study was twofold: To describe the language of clinicians and clients in a gMET intervention using a standardized MI coding system, and to use those variables to predict safer-sex outcomes three months later. Although both aims were accomplished, the primary achievement of this study was the former.

Description. Having adapted an MI sequential-coding system to groups and incarcerated adolescents, obtained reasonable inter-rater reliability, offered an initial description of MI-related behaviors within a gMET intervention, and applied those findings to the fidelity of a randomized controlled trial provides a substantial contribution to the scant literature about the process of group-delivered MI-based interventions. Although having coded fewer sessions than anticipated perhaps limited the generalizability of these findings, the project nonetheless yielded estimates of expected total utterances, behavior-count frequencies, and group-specific global ratings, as well as collected tens of thousands of transition probabilities that can be applied to sequential hypotheses. Previous studies have gathered these detailed data for individual MI sessions

(e.g., Amrhein et al., 2003; Gaume et al., 2010; Moyers et al., 2009; Vader et al., 2010), and by doing so have greatly advanced MI process research. For groups, such data are currently novel, and will be in high demand as researchers increasingly seek to translate individual MI to group settings with evidence-based guidance.

Prediction. The finding that group-level change talk was not predictive of client outcomes is disappointing not because the proposed models were incorrect—after all, falsification is the purpose of science—but because it is impossible to know why. As with all null-hypothesis significance tests, the inability here to predict safer-sex variables from group-level change talk indicated either that there was no effect present, or that it was simply not detected through these predictors and this methodology. That does not mean that it is not detectable through other methods, or present in a different sample, or even attainable by a different team of coders.

With those caveats in mind, the following are some of the most likely explanations for the results, based upon the data: (1) the components of power (e.g., inadequate sample size, tiny effect size, excessive error variance) were insufficient to detect the effect, (2) change talk about safer sex is not relevant to safer-sex target behaviors, (3) group MI-based interventions operate by different mechanisms than individual therapies, and change talk is not a mechanism of gMET, (4) group-level change talk serves as an inappropriate proxy for individual-level readiness or intention to change, (5) the content and milieu of the gMET intervention are too different from MI to be captured adequately using an MI coding system, or (6) the MISC 2.5 is inappropriate for coding groups and/or adolescents.

First, low statistical power cannot be overlooked as a likely explanatory factor. Due to missing coding, interventions with one participant, and many participants not reporting sexual intercourse between visits, many fewer participants than anticipated were included in the final samples for statistical analysis. With only $n = 45$ remaining participants, this study was unlikely to have been adequately powered even for individual-level analyses, let alone for group-level analyses, which were reduced to $n = 34$ after exclusions.

Second, the correlation matrices provided limited evidence that change talk was related to safer-sex variables through the significant but small relation ($r = 0.399$) between Percentage Change Talk and Safer Sex–Baseline–Group. Because the Baseline intervention (and therefore change talk) occurred after the safer-sex behavior, and change talk was not related to safer-sex outcome variables after the intervention, change talk within the intervention might only be interpretable as a possible marker of initial readiness to change safer-sex behavior; this conceptualization echoes Apodaca and Longabaugh's (2009) idea that client language serves as a signal of current intention to change (or not). Notably, the Risky Sex Index outcome variable had been intended to be calculated over the past three months in order to represent sexual risk between Baseline and Three Months, but that item was inadvertently excluded from the test battery for most participants in Project MARS. Instead, the Risky Sex Index was formed from a question about the participant's lifetime sexual activity and not only for the previous three month; perhaps such a variable would not necessarily be sensitive to small temporal changes, and might not have related directly to current change talk.

Third, it cannot be determined from these analyses whether gMET operates due to different factors than individual MI. However, when sequential analyses are performed, they might help to explain how clinicians and clients interact, which could then be related to sequential data from individual MI sessions (e.g., Gaume et al., 2008; Moyers et al., 2009), which have known behavioral correlates. Although it would be impossible to ascertain transitions between clients in this study because change talk was conceptualized at the group level, transitions between different clients could be captured in another study.

Fourth, the question of whether change talk is best captured at the individual or group level in group interventions also cannot be answered based upon this study. Observed standard deviations in group-level change talk were large, which suggests that using group averages to predict individual-level outcomes might have obscured important differences between change talk and outcomes. However, a methodology that allowed for the capture of individual-level speech contributions within groups could provide both individual- and group-level change talk with little additional labor. That approach might offer the most flexibility in planning coding for future MI interventions with a group-process focus, and help more easily answer questions at both the individual and group levels.

Finally, the gMET groups in this study represented an intentionally blended intervention of an MI style with educational and feedback-based components. Such interventions have received mixed support in previous studies, but several have positively influenced client outcomes (e.g., LaChance et al., 2009; Sobell & Sobell, 2009). The behavior counts and global ratings for MI fidelity from the MISC 2.5, which hung just at

the cutoff points for most variables, were consistent with an intervention that shares some characteristics of MI but is not purely MI. In that way, the measure appeared to capture most, but not all, aspects of the intervention; for example, coders noted that moments of clinician teaching in the group felt different than either information-giving or advice without permission, and perhaps alternate codes would provide a better fit for gMET. Importantly, delivering MI interventions in groups also increases the complexity required of clinicians, particularly in optimally balancing the changing needs of different group members (e.g., Feldstein Ewing et al., 2012). It is possible that existing MI fidelity measures remain appropriate for groups, but that cutoffs should be higher to account for such complexity.

Strengths and Innovations

This study showed many strengths. First, the parent study, Project MARS, was well-conceived and meticulously organized, which inspired confidence in the quality of questionnaire and in-session data. Second, this study benefitted greatly from the use of technologically advanced coding software and validated coding systems derived from Project ELICIT and other studies through UNM CASAA, as well as the expert oversight of coder training and supervision. Third, it attempted to connect in-session behaviors to post-session outcomes, which, despite null results, potentially increased our understanding of the mechanisms of group-delivered therapy. Finally, it maximized scientific benefits by making efficient use of existing data through secondary analysis.

This study made several novel contributions to the literature: (1) It was the first to relate in-group processes to safer-sex behavior, (2) it was the second to investigate in-session clinician language and client change talk in groups and relate them to outcomes,

(3) it was among the earliest to investigate the relation between a hypothesized mechanism of MI (change talk) and distal outcomes within a group setting.

Limitations

Several limitations should be noted. First, as mentioned previously, the total number of participants and groups was smaller than expected; because the number of individuals described in the initial power analysis were estimated based upon standard effect sizes and not known ones from the literature, the study almost certainly was underpowered to detect anything but a very large effect size.

Second, inter-rater reliability was favorable overall, but there was inconsistency between the reliability of the Percentage Change Talk summary variable ("excellent") and its component variables ("fair" and "good"). That discrepancy lent itself to several possible interpretations, including that (1) the ICC for Percentage Change Talk was inflated simply because summary variables tend to improve inter-rater reliability, and thus coder agreement was overestimated, (2) good reliability of Sustain Talk, despite its relative infrequency, helped compensate for poorer reliability on Change Talk, (3) the inclusion of three training recordings in the reliability sample meant coders had not yet calibrated to the revised coding system, or (4) coders simply disagreed about how utterances should be characterized. The perils of inconsistent inter-rater reliability for Percentage Change Talk are that confidence in the validity of the data is threatened, which potentially limits interpretation and generalization to other samples, and that Type II error rate increases (Hallgren, 2012), which means important relations with other variables might go undetected. However, exploratory restriction of the reliability sample to only the non-training sample greatly improved reliability of the component variables

for Percentage Change Talk and for most other summary variables. Because those sessions were coded after initial training, it is possible that they better reflect the true reliability of the sample and therefore that coders became more skilled over time, although other factors could be responsible as well. Unfortunately, it is unknowable whether coders would have become more reliable with additional practice and reliability checks, or whether they actually might have drifted over time.

Third, group participants were not recorded with individual microphones, so individual client speech could not be captured reliably or related to individual outcomes. If change talk at the group level is unimportant in predicting outcomes and only what individuals say matter, then the current methodology was unsuitable for describing the operation of MET groups. However, the results did not address whether the participation *between* group members might have elevated the shared experience of the group above what would be experienced by members of a hub-and-spokes group (i.e., a clinician interacting only with each client individually). Furthermore, this study specifically addressed whether group-level effects were related to group-level outcomes instead of individual ones (i.e., through hypotheses 1b and 2b) and that idea was not supported by the data, although small sample size was problematic.

Fourth, the number of clinicians was insufficient to allow for extensive study of between-clinician differences. Moreover, clinicians were inherently confounded with groups, so clinician effects could not be easily disentangled. Although clinician differences in delivering the gMET intervention were minimized through manualized procedures, minimum-proficiency standards, and regular supervision meetings, it would have been a helpful to be able to rule out alternate explanations for the discrepancy

between low and high global scores (e.g., characteristics of the clinician, training, intervention, particular grouping of clients, or even coder).

Fifth, only self-report was used to determine client real-world safer-sex outcomes, and it is possible that detained adolescents incorrectly reported their use of condoms and other behaviors—particularly if sex occurred after they had been using alcohol or marijuana. There is evidence that self-report questionnaire data can be reliable up to three to six months later (Napper, Fisher, Reynolds, & Johnson, 2010; Sieving et al., 2005). However, because sexual behavior certainly is inherently private and potentially stigmatizing, collecting such sensitive information accurately is viewed as a challenge of conducting research in this field, although computerized assessments are believed to provide a safe atmosphere for providing sexual information and can decrease measurement error (e.g., Schroeder, Carey, & Vanable, 2003).

Sixth, the group included mixed content from motivational, risk-reduction, and educational interventions. Unfortunately, it is impossible to determine whether a particular component of the blended intervention (e.g., feedback, education, or MI "spirit" and techniques) is responsible for any differences in outcomes, except by breaking down the intervention into those individual hypothesized components and relating them to outcomes, as we have done here. Furthermore, based upon previous work (e.g., Schmiege et al., 2009), it might have been unrealistic to expect that MET would be directly responsible for outcomes, because in that study the addition of MET to the active-treatment protocol enhanced only mediators, and not outcome efficacy.

Finally, the MISC 2.5 coding system was developed for coding individual adult MI sessions with substance-use concerns, and prior to this study not been tested with

either groups, adolescents, or safer-sex behaviors. One empirical study (e.g., Baer et al., 2008) suggests that adolescents do speak differently about change than do adults (i.e., using more Desire and Ability change language, and less Commitment), which is similar to anecdotal evidence from our coding lab (e.g., that very young adult clients tend to speak in the second person to describe their own intentions to change), but again, those ideas remain untested.

Context within the Existing Literature

In relating the current findings to the larger context of MI and MET groups, direct comparisons are severely limited by the lack of previous literature. As introduced earlier, studies of individual MI have supported change talk as the important middle link of the causal chain between clinician behaviors and client substance-use outcomes (e.g., Moyers et al., 2009; Vader et al., 2010). A handful of other studies have linked safer-sex interventions to improved outcomes for MI with individuals (e.g., Ingersoll et al., 2005; LaBrie et al., 2008; Naar-King et al., 2006) and with groups (e.g., Bryan et al., 2009; Carey et al., 1997; Kalichman et al., 1999; Rosengard et al., 2007; Schmiege et al., 2009).

However, the current study was unable to establish change talk as a predictor of safer-sex outcomes, and thus the findings stand in contrast to significant results from the only study to have linked group-level change talk with improved client outcomes (Engle et al., 2010). Notably, the two studies do differ in target behavior (i.e., safer sex vs. marijuana), number of sessions (i.e., one vs. ten), and conceptualization of change talk (i.e., strength of commitment language vs. relative frequency of all change talk). Perhaps commitment language tends to occur at greater levels of readiness and over greater group

time, which is not immediate in ambivalent adolescents within only a single three-hour intervention, and is responsible for improved outcomes in the Engle et al. study.

Other content and process variables. Given the variety of strategies employed within MET interventions and the variable outcomes achieved, perhaps some intervention elements are more likely than others to actuate the active ingredients of the treatment. If so, choosing different strategies for change-talk elicitation might strengthen the MI intervention and improve the testability of change talk as an active ingredient. For example, adding a co-clinician to the existing Project MARS intervention might reduce the interpersonal complexity and allow greater focus upon core MI relational factors and specific skills. As another example, Project MARS included normative feedback about safer-sex behaviors, which has received some support in past studies (e.g., LaBrie et al., 2006; Sobell & Sobell, 2011). Increased personalization of that information (e.g., Larimer & Cronce, 2002) might enhance the relevance to individuals, and perhaps further strengthen the intervention as well by heightening discrepancy.

If change talk is unsupported in future studies as a possible active ingredient of group MI, it will be important to consider other process and content variables that might be responsible for the efficacy of many MI groups. Some clear candidates are already hypothesized within MI, such as the technical factor of suppression of sustain talk (e.g., Apodaca, 2012), and the clinician relational factors (Miller & Rose, 2009) of Rogerian accurate empathy (Moyers & Miller, 2012) and MI spirit (McCambridge et al., 2011; Moyers et al., 2005). Similarly, relational factors *between* clients have been unstudied in group MI, but hypothesized group variables from approaches tangential to MI might provide insight. Those have included motivation and insight about self-change, sparked

by the group (Sobell & Sobell, 2011); extension of in-group factors beyond the session, such as hope (e.g., Yalom); or indirect mediators of behavior, such as self-efficacy, attitudes, and intentions (LaChance et al., 2009; Schmiege et al., 2009). Finally, the context of the intervention could change the relational variables most responsible for promoting change. For example, group cohesion (e.g., Feldstein Ewing, 2012) might have been relatively low within this single-session intervention, but could become especially relevant in multi-session interventions.

Future Directions

Analyses with existing data. This dissertation study was designed so that additional analyses could be performed later using existing coding and questionnaire data. Those could include: (1) Relating clinician global ratings to client three-month outcomes, (2) Re-analyzing data using hierarchical linear modeling to try to improve model fit, (3) Using sequential client and clinician exchange data ("transition probabilities") obtained through MISC 2.5 coding to calculate conditional probabilities of client behaviors in groups based upon clinician behaviors, via GSEQ software (Bakeman & Quera, 1995), (4) Assessing the influence of clinician speech and global MI ratings upon longer-term client outcomes after Project MARS is complete, (5) Determining whether "removed change talk" by adolescents about other group members or hypothetical adolescents serves a similar function to change talk in adults, (6) Considering whether change talk is related to condom-use intentions and attitudes (as in Schmiege et al., 2009) in adolescents who still might be contemplating behavior change, and (7) Evaluating the trajectory of group change language throughout the session by rating the strength of client change talk throughout the session (using the MISC 2.1

coding system; Miller, Moyers, Ernst, & Amrhein, 2008), and then relating it to both clinician behaviors and client outcomes.

Later steps. Ideally, the 17 sessions that were parsed but not sequentially coded could be completed by a member of the original coding team using either remaining funding resources or volunteer labor prior to those students' departure from UNM. However, securing supplemental funding for coder training, coder supervision, and a small local coding team could also facilitate parsing and coding of any Project MARS sessions created after the start of this study; that would both increase the power to detect effects of change talk upon safer-sex outcomes, and also jumpstart a local coding group-lab that could be sustained through future projects.

More distant research directions could include submission of an R21, R01, or similar NIH grant to explore the mediational chain of clinician behaviors to client change talk to client outcomes in a group setting, if client and clinician behaviors first can be related to outcomes. This study generated pilot data to determine the value and feasibility of such a project, which could be used later to lend credibility to a federal grant application. Large-scale funding for such a project could allow for significant future exploration of in-group therapeutic processes of gMET, and could form the basis of a post-graduate line of research.

If the relation between change talk and outcomes can be established in a gMET setting, it will be important to continue with this area of study to determine the sequence of events and the likely reasons for that relation—specifically, whether clinicians or other clients can influence the change talk and/or eventual behavioral change of clients in a group setting (as clinicians seem to be able to in an individual therapy setting), and

whether any other variable might better explain the relation between change talk and outcomes. Those advances would help MI to become an even more efficient treatment through evidence-based delivery in a group setting.

Figures

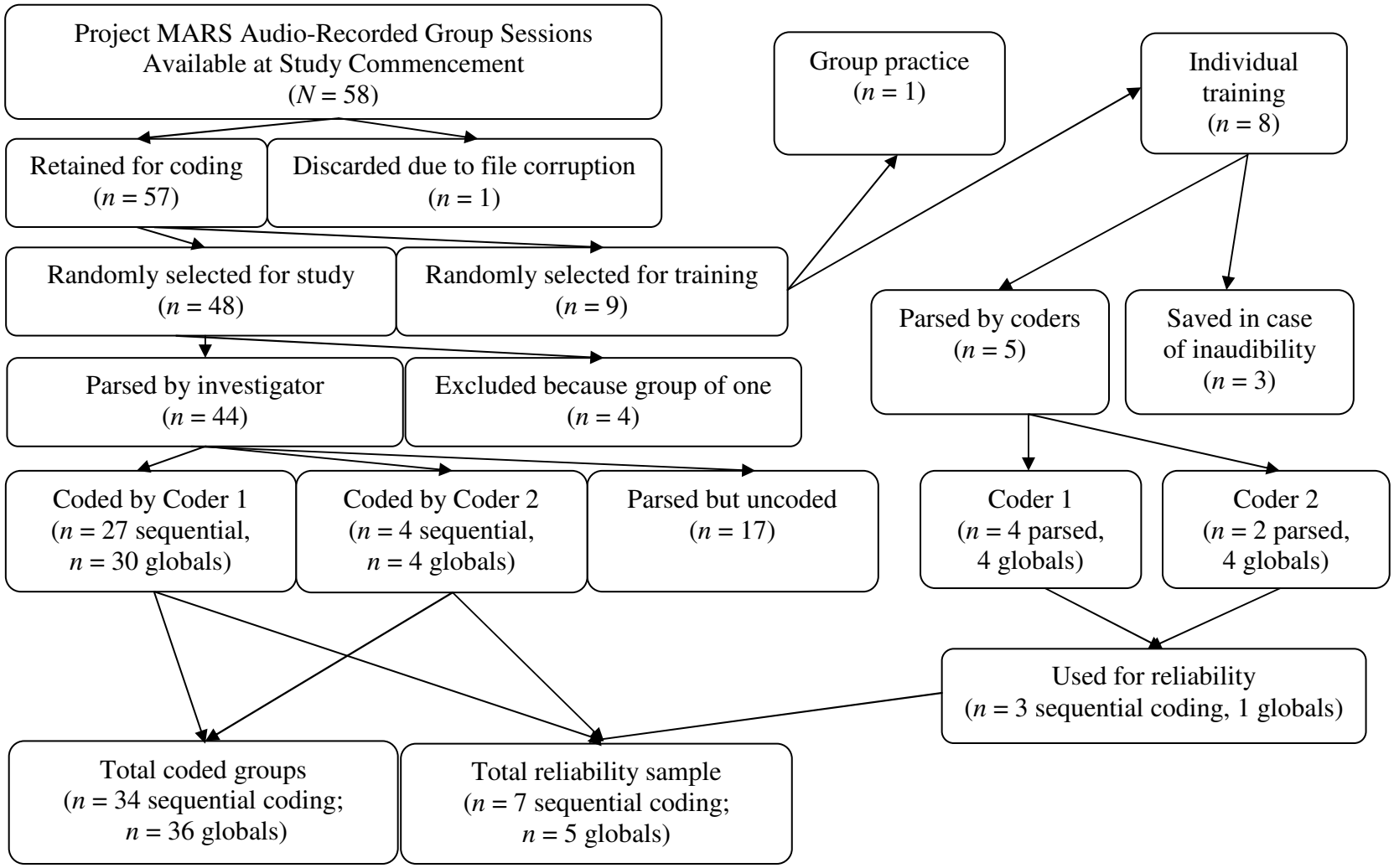

Figure 1. Flowchart of session selection and reasons for exclusion from analyses.

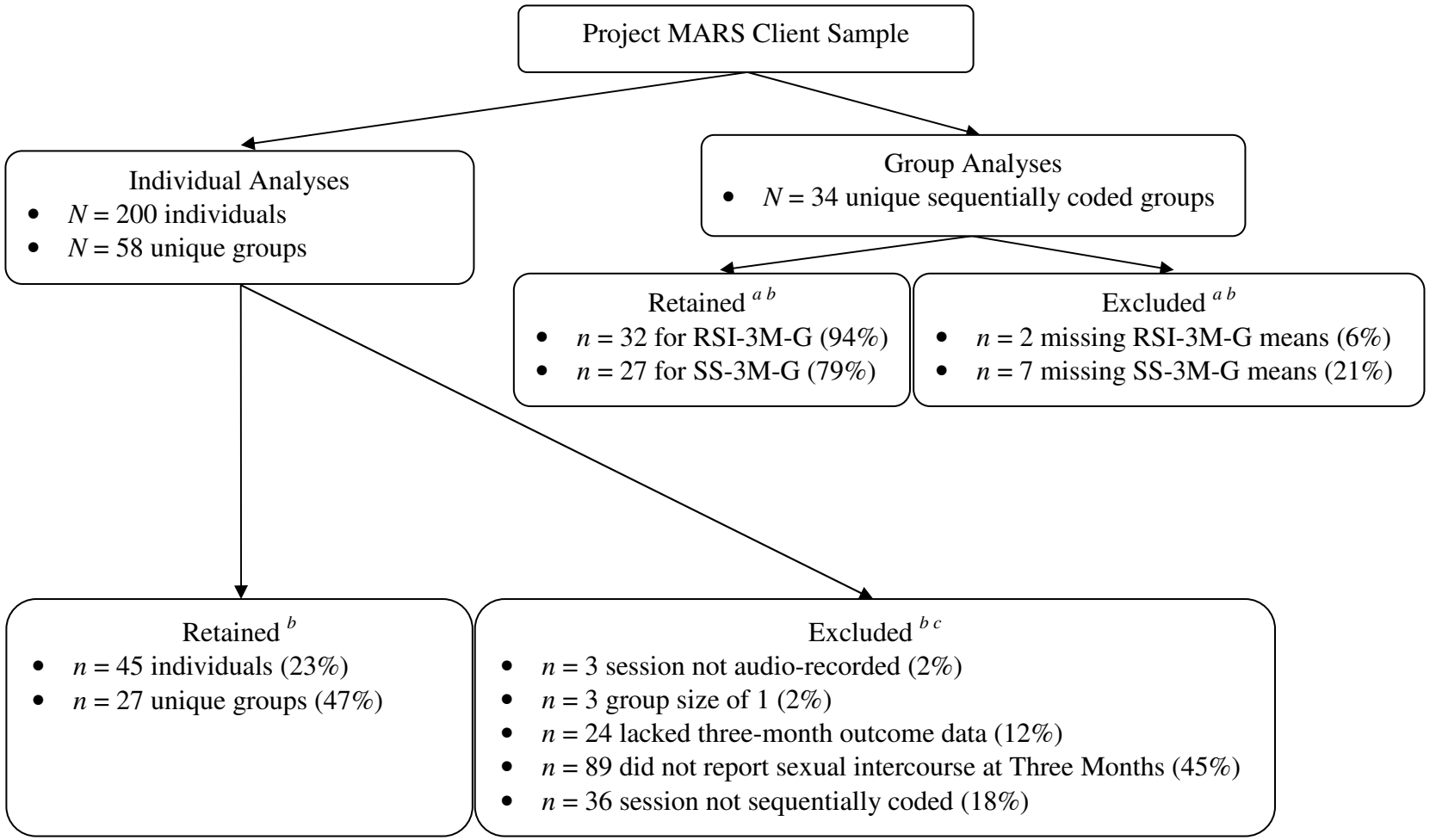

Figure 2. Flowchart of participants and reasons for exclusion from individual and group analyses.
[a] RSI-3M-G = Risky Sex Index–Three Month–Group; SS-3M-G = Safer Sex–Three Month–Group. [b] Percentages are of original *N*s, and are rounded. [c] Participants could be excluded for multiple reasons, but only one was listed in attrition analyses.

Tables

Table 1. Key Variables

Variable Name (Abbreviation)	Time [a]	How Formed [b]	Predicts [c]
Grouping Variables			
Intercourse	3M	Item t2j25 (only for individual analyses); dichotomized into yes/no within the past three months	• N/A
Group	BL	Filter applied to select only one case per group for selected/not selected (only for group analyses)	• N/A
Predictor Variables			
Percentage Change Talk (% CT)	BL	Group-level variable; Change Talk / (Change Talk + Sustain Talk)	• RSI-3M-I • RSI-3M-G • SS-3M-I • SS-3M-G
Risky Sex Index–Baseline–Individual (RSI-BL-I)	BL	Items t0j06 * t0j07; for individuals	• RSI-3M-I
Risky Sex Index–Baseline–Group (RSI-BL-G)	BL	Items t0j06 * t0j07; averaged within intervention groups	• RSI-3M-G
Safer Sex–Baseline–Individual (SS-BL-I)	BL	(t0j27 + t0j28) / 2; for individuals	• SS-3M-I
Safer Sex–Baseline–Group (SS-BL-G)	BL	(t0j27 + t0j28) / 2; averaged within intervention groups	• SS-3M-G
Client Sex	BL	Two levels: 0 = girl, 1 = boy	• RSI-3M-I • RSI-3M-G • SS-3M-I • SS-3M-G
Clinician	BL	Six unweighted effect-coded variables	• RSI-3M-I • RSI-3M-G • SS-3M-I • SS-3M-G
Outcome Variables			
Risky Sex–Index Three Month–Individual (RSI-3M-I)	3M	Items t2j06 * t2j07; for individuals	• N/A
Risky Sex Index–Three Month–Group (RSI-3M-G)	3M	Items t2j06 * t2j07; averaged within intervention groups	• N/A
Safer Sex–Three Month–Individual (SS-3M-I)	3M	(t2j27 + t2j28) / 2; for individuals	• N/A

Safer Sex–Three Month–Group (SS-3M-G)	3M	(t2j27 + t2j28) / 2; averaged within intervention groups	• N/A

[a] BL = Baseline; 3M = Three Month. [b] Item names refer to original variables in the study manual (Project MARS, n.d.). [c] RSI = Risky Sex Index; SS = Safer Sex; I = individual; G = group.

Table 2. Risky Sex Index–Three-Month–Individual as Dependent Variable

Variable [a]	*b*	*SE*	*t*	*p*
Constant	-1.545	7.103	-0.217	.829
Client Sex	1.552	1.054	1.472	.149
% CT	8.158	8.983	0.908	.369
RSI-BL-I	0.422	0.170	2.476	.018*

[a] % CT = Percentage Change Talk; RSI-BL-I = Risky Sex Index–Baseline–Individual.
* $p < 0.05$, two-tailed.

Table 3. Risky Sex Index–Three-Month–Individual Correlation Matrix

Predictor [a]	Client Sex	% CT	RSI-BL-I
Client Sex	-	-	-
% CT	0.008	-	-
RSI-BL-I	0.022	-0.019	-
RSI-3M-I	0.217	0.133	0.355*

[a] % CT = Percentage Change Talk; RSI-BL-I = Risky Sex Index–Baseline–Individual; RSI-3M-I = Risky Sex Index–Three Month–Individual.
* $p < 0.05$, two-tailed.

Table 4. Risky Sex Index–Three Month–Group as Dependent Variable

Variable [a]	*b*	*SE*	*t*	*p*
Constant	-2.936	8.366	-0.351	.728
Client Sex	0.891	0.958	0.930	.360
% CT	12.012	9.957	1.206	.238
RSI-BL-G	0.308	0.232	1.328	.195

[a] % CT = Percentage Change Talk; RSI-BL-G = Risky Sex Index–Baseline–Group.

Table 5. Risky Sex Index–Three Month–Group Correlation Matrix

Predictor [a]	Client Sex	% CT	RSI-BL-G
Client Sex	-	-	-
% CT	-0.057	-	-
RSI-BL-G	-0.225	-0.102	-
RSI-3M-G	0.086	0.171	0.181

[a] % CT = Percentage Change Talk; RSI-BL-G = Risky Sex Index–Baseline–Group; RSI-3M-G = Risky Sex Index–Three Month–Group.

Table 6. Safer Sex–Three Month–Individual as Dependent Variable

Variable [a]	*b*	*SE*	*t*	*p*
Constant	-0.162	1.417	-0.114	.910
Client Sex	0.406	0.213	1.902	.066
% CT	0.967	1.865	0.518	.608
SS-BL-I	0.658	0.175	3.770	.001**

[a] % CT = Percentage Change Talk; SS-BL-I = Safer Sex–Baseline–Individual.
** $p < 0.01$, two-tailed.

Table 7. Safer Sex–Three Month–Individual Correlation Matrix

Predictor [a]	Client Sex	% CT	SS-BL-I
Client Sex	-	-	-
% CT	0.008	-	-
SS-BL-I	0.015	0.170	-
SS-3M-I	0.283	0.204	0.537**

[a] % CT = Percentage Change Talk; SS-BL-I = Safer Sex–Baseline–Individual; SS-3M-I = Safer Sex–Three Month–Individual.
** $p < 0.01$, two-tailed.

Table 8. Safer Sex–Three Month–Group as Dependent Variable

Variable [a]	*b*	*SE*	*t*	*p*
Constant	-1.150	1.685	-0.682	.502
Client Sex	0.444	0.230	1.931	.066
% CT	1.296	2.349	0.552	.586
SS-BL-G	0.915	0.401	2.280	.032*

[a] % CT = Percentage Change Talk; SS-BL-G = Safer Sex–Baseline–Group.
* $p < 0.05$, two-tailed.

Table 9. Safer Sex–Three Month–Group Correlation Matrix

Predictor [a]	Client Sex	% CT	SS-BL-G
Client Sex	-	-	-
% CT	-0.057	-	-
SS-BL-G	-0.071	0.399*	-
SS-3M-G	0.302	0.260	0.442*

[a] % CT = Percentage Change Talk; SS-BL-G = Safer Sex–Baseline–Group; SS-3M-G = Safer Sex–Three Month–Group.
* $p < 0.05$, two-tailed.

Table 10. Coder Reliability for Behavior-Count Summary Variables (Total Reliability Sample and Non-Training Sample)

Summary Variable	ICC (Total)	Range (Total) [a b]	ICC (Non-Training)	Range (Non-Training) [b c]
Change Talk	0.435	Fair	0.986	Excellent
Sustain Talk	0.701	Good	0.863	Excellent
Percentage Change Talk	0.795	Excellent	0.690	Good
Percentage Open Questions	0.864	Excellent	0.981	Excellent
Reflection to Question Ratio	0.716	Good	0.930	Excellent
Percentage Reflections	0.780	Excellent	0.930	Excellent
MI Consistent	0.864	Excellent	0.959	Excellent
MI Inconsistent	0.380	Poor	-0.117	Uninterpretable
Percentage MI Consistent	0.472	Fair	0.016	Uninterpretable
Closed Questions	0.936	Excellent	0.985	Excellent
Open Questions	0.937	Excellent	0.987	Excellent
Giving Information	0.516	Fair	0.631	Good
Follow/Neutral	0.899	Excellent	0.982	Excellent

[a] Original reliability sample (n = 7). [b] Cicchetti (1994). [c] Non-training reliability sample (n = 4).

Table 11. Summary of Selected Behavior-Count Summary Variables

Summary Variable	*M*	*SD*	Basic Competence/ Proficiency? [a]
Clinician Behaviors			
Giving Information	149.53	30.71	-
Closed Question	113.21	38.47	-
Open Question	78.97	30.97	-
Simple Reflection			
Change Talk	33.76	14.39	-
Sustain Talk	8.26	4.86	-
Neutral	81.29	34.40	-
Complex Reflection			
Change Talk	4.18	3.16	-
Sustain Talk	0.91	1.31	-
Neutral	13.21	8.22	-
MI Consistent	296.82	85.03	-
MI Inconsistent	8.38	3.54	-
Percentage MI Consistent	97.03%	1.52%	Yes/No
Percentage Open Questions	40.95%	11.22%	Yes/No
Percentage Reflections	42.78%	6.97%	-
Percentage Complex Reflections	13.21%	7.13%	No/No
Reflection to Question Ratio	77.22%	21.04%	No/No
Client Behaviors			
Change Talk	62.82	23.62	-

Sustain Talk	19.12	8.08	-
Percentage Change Talk	76.28%	9.01%	-
Follow/Neutral	286.53	67.44	-
No Code	28.15	13.61	-

[a] Per MITI 3.1.2 standards (Moyers et al., 2012).

Table 12. Summary of Global Ratings

Summary Variable	*M*	*SD*	Minimum	Maximum	Basic Competence/ Proficiency? [a]
Acceptance	4.05	0.33	2	4	Yes/Yes
Empathy	3.62	0.72	2	4	Yes/Yes
Direction	5.00	0.00	5	5	Yes/Yes
Autonomy Support	3.43	0.60	2	4	No/No
Collaboration	3.49	0.56	2	4	No/No
Evocation	3.41	0.69	2	5	No/No
MI Spirit	3.44	0.41	2.33	4.33	No/No
Self-Exploration	2.39	0.55	2	4	-

[a] Per MITI 3.1.2 standards (Moyers et al., 2012).

References

Amrhein, P. C., Miller, W. R., Yahne, C. E., Palmer, M., & Fulcher, L. (2003). Client commitment language during motivational interviewing predicts drug use outcomes. *Journal of Consulting and Clinical Psychology, 71*(5), 862-878.

Apodaca, T. R. (2012, June). Sustain talk in MI predicts drinking outcomes, while change talk does not. Retrieved June 3, 2013, from http://www.icmi2012. com/ wp-content/uploads/2012/10/apodaca-sustain-talk.pdf

Apodaca, T. R., & Longabaugh, R. (2009). Mechanisms of change in motivational interviewing: A review and preliminary evaluation of the evidence. *Addiction, 104*(5), 705-715.

Baer, J. S., Beadnell, B., Garrett, S. B., Hartzler, B., Wells, E. A., & Peterson, P. L. (2008). Adolescent change language within a brief motivational intervention and substance use outcomes. *Psychology of Addictive Behaviors, 22*(4), 570-575.

Bailey, K. A., Baker, A. L., Webster, R. A., & Lewin, T. J. (2004). Pilot randomized controlled trial of a brief alcohol intervention group for adolescents. *Drug and Alcohol Review, 23*(2), 157-166.

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

Barnet, B., Liu, J., DeVoe, M., Duggan, A. K., Gold, M. A., & Pecukonis, E. (2009). Motivational intervention to reduce rapid subsequent births to adolescent mothers: a community-based randomized trial. *Annals of Family Medicine, 7*(5), 436-445.

Brafford, L. J., & Beck, K.H. (1991). Development and validation of a condom self-efficacy scale for college students. *Journal of American College Health, 39*(5), 219-225.

Brien, T.M., Thombs, D. L., Mahoney, C. A., & Wallanu, L. (1994). Dimensions of self-efficacy among three distinct groups of condom users. *Journal of American College Health, 42*, 167-174.

Bryan, A., Aiken, L. S., & West, S. G. (2004). HIV/STD risk among incarcerated adolescents: Optimism about the future and self-esteem as predictors of condom use self-efficacy. *Journal of Applied Social Psychology, 34*(5), 912-936.

Bryan, A., Rocheleau, C. A., Robbins, R. N., & Hutchison, K. E. (2005). Condom use among high-risk adolescents: Testing the influence of alcohol use on the relationship of cognitive correlates of behavior. *Health Psychology, 24*(2), 133-142.

Bryan, A. D., Schmiege, S. J., & Broaddus, M. R. (2009). HIV risk reduction among detained adolescents: A randomized, controlled trial. Pediatrics, 124, e1180-e1188. doi: 10.1542/peds.2009-0679

Burke, B. L., Arkowitz, H., & Menchola, M. (2003). The efficacy of motivational interviewing: a meta-analysis of controlled clinical trials. *Journal of Consulting and Clinical Psychology, 71*(5), 843-861.

Burleson, J. A., Kaminer, Y., & Dennis, M. L. (2006). Absence of iatrogenic or contagion effects in adolescent group therapy: Findings from the Cannabis Youth Treatment (CYT) study. *American Journal on Addictions, 15*(S1), 4-15.

Campbell, S. D., Adamson, S. J., & Carter, J. D. (2010). Client language during Motivational Enhancement Therapy and alcohol use outcome. *Behavioural and Cognitive Psychotherapy, 38*(4), 399-415. doi: 10.1017/S1352465810000263

Canterbury, R. J., Clavet, G. J., McGarvey, E. L., & Koopman, C. (1998). HIV risk-related attitudes and behaviors of incarcerated adolescents: Implications for public school students. *The High School Journal, 82*(1), 1-10. doi: 10.2307/40364705

Carey, M. P., Braaten, L. S., Maisto, S. A., Gleason, J. R., Forsyth, A. D., Durant, L. E., et al. (2000). Using information, motivational enhancement, and skills training to reduce the risk of HIV infection for low-income urban women: a second randomized clinical trial. *Health Psychology, 19*(1S), 3-11.

Carey, M. P., Maisto, S. A., Kalichman, S. C., Forsyth, A. D., Wright, E. M., & Johnson, B. T. (1997). Enhancing motivation to reduce the risk of HIV infection for economically disadvantaged urban women. *Journal of Consulting and Clinical Psychology, 65*(4), 531-541.

Carroll, K. M., Ball, S. A., Nich, C., Martino, S., Frankforter, T. L., Farentinos, C., et al. (2006). Motivational interviewing to improve treatment engagement and outcome in individuals seeking treatment for substance abuse: a multisite effectiveness study. *Drug Alcohol Depend, 81*(3), 301-312.

Catley, D., Harris, K. J., Mayo, M. S., Hall, S., Okuyemi, K. S., Boardman, T., et al. (2006). Adherence to principles of motivational interviewing and client within-session behavior. *Behavioural and Cognitive Psychotherapy, 34*(1), 43-56.

Center for Substance Abuse Treatment (CSAT). (1999). *Enhancing motivation for change in substance abuse treatment.* (Treatment Improvement Protocol Series

35). Rockville, MD: Substance Abuse and Mental Health Services Administration.

Center for Substance Abuse Treatment (CSAT). (2005). *Quick guide for clinicians based on TIP 41 substance abuse treatment: Group therapy*. Treatment Improvement Protocol (TIP) 41. DHHS Publication No SMA (05-3991). Rockville, MD: Substance Abuse and Mental Health Services Administration.

Center for Substance Abuse Treatment (CSAT). (2006). *Substance abuse: Clinical issues in intensive outpatient treatment* (Treatment Improvement Protocol Series 47). Rockville, MD: Substance Abuse and Mental Health Services Administration.

Chen, X., Murphy, D. A., Naar-King, S., & Parsons, J. T. (in press). A clinic-based motivational intervention improves condom use among subgroups of youth living with HIV: A multicenter randomized controlled trial. *Journal of Adolescent Health*. doi: 10.1016/j.jadohealth.2010.11.252

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychological Assessment, 6, 284–290.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155-159.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika. 16*, 297-334.

D'Amico, E. J., Ewing, S. W. F., Engle, B., Hunter, S., Osilla, K. C., & Bryan, A. (2011). Group alcohol and drug treatment. In S. Naar-King (Ed.), *Motivational interviewing with adolescents and young adults* (pp. 151-157). New York: Guilford.

D'Amico, E. J., Osilla, K. C., & Hunter, S. B. (2010). Developing a group motivational interviewing intervention for first-time adolescent offenders at-risk for an alcohol or drug use disorder. *Alcoholism Treatment Quarterly, 28*(4), 417 - 436.

D'Amico, E. J., Osilla, K. C., Miles, J. N. V., Ewing, B., Sullivan, K., Katz, K., & Hunter, S. B. (2012). Assessing motivational interviewing integrity for group interventions with adolescents. *Psychology of Addictive Behaviors*. doi: 10.1037/a0027987

Dimeff, L. A., & McNeely, M. (2000). Computer-enhanced primary care practitioner advice for high-risk college drinkers in a student primary health-care setting. *Cognitive and Behavioural Practice, 7*(1), 82–100.

Dishion, T. J., & Dodge, K. A. (2005). Peer contagion in interventions for children and adolescents: Moving toward an understanding of the ecology and dynamics of change. *Journal of Abnormal Child Psychology, 33*(3), 395-400.

Donovan, D. M., Rosengren, D. B., Downey, L., Cox, G. B., & Sloan, K. L. (2001). Attrition prevention with individuals awaiting publicly funded drug treatment. *Addiction, 96*(8), 1149-1160.

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

Engle, B., Macgowan, M. J., Wagner, E. F., & Amrhein, P. C. (2010). Markers of marijuana use outcomes within adolescent substance abuse group treatment. *Research on Social Work Practice, 20*(3), 271-282. doi: 10.1177/1049731509347855

Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior Research Methods, Instruments, & Computers*, *28*, 1-11.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191.

Feldstein Ewing, S., Walters, S. T., & Baer, J. S. (2012). Motivational interviewing groups for adolescents and emerging adults. In C. Wagner & K. Ingersoll (Eds.), *Motivational Interviewing in Groups* (pp. 387–406). New York: Guilford Press.

French, M. T., Zavala, S. K., McCollister, K. E., Waldron, H. B., Turner, C. W., & Ozechowski, T. J. (2008). Cost-effectiveness analysis of four interventions for adolescents with a substance use disorder. *Journal of Substance Abuse Treatment, 34*(3), 272-281.

Gaume, J., Bertholet, N., Faouzi, M., Gmel, G., & Daeppen, J.-B. (2010). Counselor motivational interviewing skills and young adult change talk articulation during brief motivational interventions. *Journal of Substance Abuse Treatment, 39*(3), 272-281.

Gaume, J., Gmel, G., Faouzi, M., & Daeppen, J.-B. (2008). Counsellor behaviours and patient language during brief motivational interventions: A sequential analysis of speech. *Addiction, 103*(11), 1793-1800.

Glynn, L. H., Hallgren, K. A., Houck, J. M., & Moyers, T. B. (2012). CACTI: Free, open-source software for coding behavioral interactions. *PLoS One 7*(7), e39740. doi: 10.1371/journal.pone. 0039740. Retrieved from http://www.plosone.org/article/info:doi/ 10.1371/journal.pone.0039740.

Glynn, L. H., & Moyers, T. B. (2009). *Manual for the Motivational Interviewing Skill Code (MISC), Version 1.1: Addendum to MISC 1.0*. Unpublished manuscript, Center on Alcoholism, Substance Abuse, and Addictions, University of New Mexico. Retrieved September 2, 2009, from Center on Alcoholism, Substance Abuse, and Addictions Web site: http://casaa.unm.edu/ download/MISC_1.1_Manual.pdf.

GoldWave, Inc. (2013). *GoldWave.* Retrieved from http://www.goldwave.com/

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology, 8*(1), 23-34.

Hettema, J., Steele, J., & Miller, W. R. (2005). Motivational interviewing. *Annual Review of Clinical Psychology, 1*, 91-111.

Hoag, M. J., & Burlingame, G. M. (1997). Child and adolescent group psychotherapy: A narrative review of effectiveness and the case for meta-analysis. *Journal of Child and Adolescent Group Therapy, 7*(2), 51-68.

Houck, J. M., Moyers, T. B., Miller, W. R., Glynn, L. H., & Hallgren, K. A. (2010). *ELICIT MISC 2.5 coding manual.* Unpublished manuscript, Department of Psychology, University of New Mexico.

Hountras, P., & Anderson, D. (1969). Counselor conditions for self-exploration of college students [Abstract]. *Personal and Guidance Journal, 48*, 45-48.

Ingersoll, K. S., Ceperich, S. D., Nettleman, M. D., Karanda, K., Brocksen, S., & Johnson, B. A. (2005). Reducing alcohol-exposed pregnancy risk in college women: initial outcomes of a clinical trial of a motivational intervention. *Journal of Substance Abuse Treatment 29*(3), 173–180.

Kalichman, S. C., Cherry, C., & Sperling, F. B. (1999). Effectiveness of a video-based motivational skills-building HIV risk reduction intervention for inner-city African American men. *Journal of Consulting and Clinical Psychology, 67*(6), 959–966.

Kaminer, Y. (2005). Challenges and opportunities for youth substance abuse group therapy: A critical review. *Addictive Behaviors, 30*(9), 1765–1774.

Kaminer, Y., Burleson, J. A., & Goldberger, R. (2002). Cognitive-behavioral coping skills and psychoeducation therapies for adolescent substance abuse. *Journal of Nervous and Mental Disease, 190*(11), 737-745.

Kazdin, A. E. (2007). Mediators and mechanisms of change in psychotherapy research. *Annual Review of Clinical Psychology, 3*(1), 1-27. doi: 10.1146/annurev.clinpsy.3.022806. 091432

Kazdin, A. E., & Nock, M. K. (2003). Delineating mechanisms of change in child and adolescent therapy: Methodological issues and research recommendations. *Journal of Child Psychology and Psychiatry, 44*(8), 1116–1129.

Kiene, S. M., & Barta, W. D. (2006). A brief individualized computer-delivered sexual risk reduction intervention increases HIV/AIDS preventive behavior. *Journal of Adolescent Health, 39*(3), 404-410.

Kurtz, R. R., & Grummon, D. L. (1972). Different approaches to the measurement of therapist empathy and their relationship to therapy outcomes. *Journal of Consulting and Clinical Psychology, 39*(1), 106-115.

LaBrie, J. W., Pedersen, E. R., Lamb, T. F., & Quinlan, T. (2006). A campus-based motivational enhancement group intervention reduces problematic drinking in freshmen male college students. *Addictive Behaviors, 32*(5), 889-901.

LaBrie, J. W., Pedersen, E. R., Thompson, A. D., & Earleywine, M. (2008). A brief decisional balance intervention increases motivation and behavior regarding condom use in high-risk heterosexual college men. *Archives of Sexual Behavior, 37*(2), 330-339.

LaBrie, J. W., Thompson, A. D., Huchting, K., Lac, A., & Buckley, K. (2007). A group motivational interviewing intervention reduces drinking and alcohol-related negative consequences in adjudicated college women. *Addictive Behaviors, 32*(11), 2549-2562.

LaChance, H., Feldstein Ewing, S. W., Bryan, A. D., & Hutchison, K. E. (2009). What makes group MET work? A randomized controlled trial of college student drinkers in mandated alcohol diversion. *Psychology of Addictive Behaviors, 23*(4), 598-612.

Larimer, M. E., & Cronce, J. M. (2002). Identification, prevention and treatment: a review of individual-focused strategies to reduce problematic alcohol consumption by college students. *Journal of Studies on Alcohol, Supplement*(14), 148-163.

Larimer, M. E., Turner, A. P., Anderson, B. K., Fader, J. S., Kilmer, J. R., Palmer, R. S., et al. (2001). Evaluating a brief alcohol intervention with fraternities. *Journal of Studies on Alcohol, 62*(3), 370-380.

Lincourt, P., Kuettel, T. J., & Bombardier, C. H. (2002). Motivational interviewing in a group setting with mandated clients: A pilot study. *Addictive Behaviors, 27*(3), 381-391.

Lopez, L. M., Tolley, E. E., Grimes, D. A., & Chen-Mok, M. (2011). Theory-based interventions for contraception (Review). *The Cochrane Library of Systematic Reviews, 3*(CD007249). doi: 10.1002/14651858.CD007249.pub3

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

McCambridge, J., Day, M., Thomas, B. A., & Strang, J. (2011). Fidelity to Motivational Interviewing and subsequent cannabis cessation among adolescents. *Addictive Behaviors, 36*(7), 749-754.

Martin, T., Moyers, T. B., Houck, J. M., Christopher, P. J., & Miller, W. R. (2005). *Motivational Interviewing Sequential Code for Observing Process Exchanges (MI-SCOPE) Coder's Manual*. Retrieved from http://casaa.unm.edu/download/scope.pdf

Michael, K. D., Curtin, L., Kirkley, D. E., Jones, D. L., & Harris, R., Jr. (2006). Group-based motivational interviewing for alcohol use among college students: An exploratory study. *Professional Psychology: Research and Practice, 37*(6), 629-634.

Miller, W. R. (Ed.) (2004). Combined Behavioral Intervention Manual: A clinical research guide for therapists treating people with alcohol abuse and dependence. *COMBINE Monograph Series, Volume 1. Combined Behavioral Intervention Manual: A Clinical Research Guide for Therapists Treating People with Alcohol Abuse and Dependence* (DHHS Publication No. NIH 04–5288). Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism.

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*(3), 455-461.

Miller, W. R., Moyers, T. B., Ernst, D., & Amrhein, P. (2008). *Manual for the Motivational Interviewing Skill Code (MISC), Version 2.1.* Retrieved April 15, 2009, from University of New Mexico Center on Alcoholism, Substance Use, and Addictions Web site: http://casaa.unm.edu/download/misc.pdf

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

Miller, W. R., & Rose, G. S. (2009). Toward a theory of motivational interviewing. *American Psychologist, 64*(6), 527-537.

Miller, W. R., & Sanchez, V. C. (1994). Motivating young adults for treatment and lifestyle change. In G. Howard & P. Nathan (Eds.), *Alcohol Use and Misuse by Young Adults* (pp. 55-81). Notre Dame, IN: University of Notre Dame 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*(4), 251–268.

Miller, W. R., Yahne, C., & Tonigan, J. S. (2003). Motivational interviewing in drug abuse services: A randomized trial. *Journal of Consulting and Clinical Psychology, 71*(4), 754–763.

Miller, W. R., Zweben, A., DiClemente, C. C., & Rychtarik, R. G. (1994). *Motivational enhancement therapy manual: A clinical research guide for therapists treating individuals with alcohol abuse and dependence. Project MATCH Monograph Series, Vol. 2. DHHS Publication No. 94-3723*. Rockville MD: NIAAA.

Moyers, T. B., & Houck, J. M. (2011). Combining motivational interviewing with cognitive-behavioral treatments for substance abuse: Lessons from the COMBINE research project. *Cognitive and Behavioral Practice, 18*, 38-45.

Moyers, T. B., & Martin, T. (2006). Therapist influence on client language during motivational interviewing sessions. *Journal of Substance Abuse Treatment, 30*(3), 245-251.

Moyers, T. B., Martin, T., Christopher, P. J., Houck, J. M., Tonigan, J. S., & Amrhein, P. C. (2007). Client language as a mediator of Motivational Interviewing efficacy: Where is the evidence? *Alcoholism: Clinical and Experimental Research, 31*(S3), 40S-47S.

Moyers, T. B., Martin, T., Houck, J. M., Christopher, P. J., & Tonigan, J. S. (2009). From in-session behaviors to drinking outcomes: A causal chain for motivational interviewing. *Journal of Consulting and Clinical Psychology, 77*(6), 1113-1124.

Moyers, T. B., Martin, T., Manuel, J. K., Hendrickson, S. M., & Miller, W. R. (2005). Assessing competence in the use of motivational interviewing. *Journal of Substance Abuse Treatment, 28*(1), 19-26.

Moyers, T. B., Martin, T., Manuel, J. K., Miller, W. R., & Ernst, D. (2012). *Revised Global Scales: Motivational Interviewing Treatment Integrity 3.1.2 (MITI 3.1.2).* University of New Mexico. Unpublished manuscript, University of New Mexico.

Moyers, T. B., & Miller, W. R. (2012, October 1). Is low therapist empathy toxic? *Psychology of Addictive Behaviors*. [epub ahead of print] doi: 10.1037/a0030274

Naar-King, S. (2011). Motivational interviewing in adolescent treatment. *The Canadian Journal of Psychiatry / La Revue canadienne de psychiatrie, 56*(11), 651-657.

Naar-King, S., Wright, K., Parsons, J. T., Frey, M., Templin, T., Lam, P., et al. (2006). Healthy choices: Motivational enhancement therapy for health risk behaviors in HIV-positive youth. *AIDS Education and Prevention, 18*(1), 1-11.

Napper, L. E., Fisher, D. G., Reynolds, G. L., & Johnson, M. E. (2010). HIV risk behavior self-report reliability at different recall periods. *AIDS and Behavior, 14*(1), 152-161.

National Institute on Drug Abuse. (2003). Group Therapy Research. Retrieved May 8, 2011, from http://archives.drugabuse.gov/meetings/grouptherapy.html

National Registry of Evidence-based Programs and Practices. (2007). *Motivational enhancement therapy*. Retrieved November 30, 2008, from http://www.nrepp.samhsa.gov/ programfulldetails.asp?PROGRAM_ID=182

O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality and Quantity, 41*(5), 673-690.

O'Leary, T. A., Brown, S. A., Colby, S. M., Cronce, J. M., D'Amico, E. J., Fader, J. S., et al. (2002). Treating adolescents together or individually? Issues in adolescent

substance abuse interventions. *Alcoholism: Clinical and Experimental Research, 26*(6), 890-899.

Patterson, G. R., & Forgatch, M. S. (1985). Therapist behavior as a determinant for client noncompliance: a paradox for the behavior modifier. *Journal of Consulting and Clinical Psychology, 53*(6), 846-851.

Project MARS. (2010). *Your Call: An HIV/STD prevention program: Intervention leader's manual.* Unpublished manuscript, University of New Mexico.

Project MARS. (n.d.). *University of New Mexico Adolescent 3Month Follow-Up Health Survey*. Unpublished manuscript, University of New Mexico.

Project MATCH Research Group. (1998). Matching patients with alcohol disorders to treatments: Clinical implications from Project MATCH. *Journal of Mental Health, 7*(6), 589-602.

Renjilian, D. A., Perri, M. G., Nezu, A. M., McKelvey, W. F., Shermer, R. L., & Anton, S. D. (2001). Individual versus group therapy for obesity: Effects of matching participants to their treatment preferences. *Journal of Consulting and Clinical Psychology, 69*(4), 717-721.

Rogers, C. R. (1965). *Client-centered therapy: Its current practice, implications, and theory.* Boston: Houghton Mifflin.

Rosengard, C., Stein, L. A. R., Barnett, N. P., Monti, P. M., Golembeske, C., & Lebeau-Craven, R. (2006). Co-occurring sexual risk and substance use behaviors among incarcerated adolescents. *Journal of Correctional Health Care, 12*(4)*,* 279–287.

Sacco, W. P., Levine, B., Reed, D. L., & Thompson, K. (1991). Attitudes about condom use as an AIDS-relevant behavior: Their factor structure and relation to condom

use. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 3*(2), 265-272.

Santa Ana, E. J., Wulfert, E., & Nietert, P. J. (2007). Efficacy of group motivational interviewing (GMI) for psychiatric inpatients with chemical dependence. *Journal of Consulting and Clinical Psychology, 75*(5), 816-822. doi: 10.1037/0022-006X.75.5.816

Schermer, C. R., Moyers, T. B., Miller, W. R., & Bloomfield, L. A. (2006). Trauma center brief interventions for alcohol disorders decrease subsequent driving under the influence arrests. *Journal of Trauma, 60*(1), 29-34.

Schmiege, S. J., Broaddus, M. R., Levin, M., & Bryan, A. D. (2009). Randomized trial of group interventions to reduce HIV/STD risk and change theoretical mediators among detained adolescents. *Journal of Consulting and Clinical Psychology, 77*(1), 38-50.

Shrier, L. A., Ancheta, R., Goodman, E., Chiou, V. M., Lyden, M. R., & Emans, S. J. (2001). Randomized controlled trial of a safer sex intervention for high-risk adolescent girls. *Archives of Pediatrics and Adolescent Medicine, 155*(1), 73-79.

Sieving, R., Hellerstedt, W., McNeely, C., Fee, R., Snyder, J., & Resnick, M. (2005). Reliability of self-reported contraceptive use and sexual behaviors among adolescent girls. *The Journal of Sex Research, 42*(2), 159-166.

Slavet, J. D., Stein, L. A. R., Klein, J. L., Colby, S. M., Barnett, N. P., & Monti, P. M. (2005). Piloting the Family Check-Up with incarcerated adolescents and their parents. *Psychological Services, 2*(2), 123-132.

Sobell, L. C., & Sobell, M. B. (2011). *Group Therapy for Substance Use Disorders: A Motivational Cognitive-Behavioral Approach*. New York: Guilford Press.

Sobell, L. C., Sobell, M. B., & Agrawal, S. (2009). Randomized controlled trial of a cognitive–behavioral motivational intervention in a group versus individual format for substance use disorders. *Psychology of Addictive Behaviors, 23*(4), 672-683. doi: 10.1037/a0016636

St. Lawrence, J. S., Reitman, D., Jefferson, K. W., Alleyne, E., Brasfield, T. L., & Shirley, A. (1994). Factor structure and validation of an adolescent version of the Condom Attitude Scale: An instrument for measuring adolescents' attitudes towards condoms. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 6*, 352–359.

Stern, S. A., Meredith, L. S., Gholson, J., Gore, P., & D'Amico, E. J. (2007). Project CHAT: A brief motivational substance abuse intervention for teens in primary care. *Journal of Substance Abuse Treatment, 32*(2), 153-165.

Strang, J., & McCambridge, J. (2004). Can the practitioner correctly predict outcome in motivational interviewing? *Journal of Substance Abuse Treatment, 27*(1), 83-88.

Tolou-Shams, M., Stewart, A., Fasciano, J., & Brown, L. K. (2010). A review of HIV prevention interventions for juvenile offenders. *Journal of Pediatric Psychology, 35*(3), 250-261. doi: 10.1093/jpepsy/jsp069

Truax, C. B. (1968). Therapist interpersonal reinforcement of client self-exploration and therapeutic outcome in group psychotherapy. *Journal of Counseling Psychology, 15*(3), 225-231.

Truax, C. B., & Carkhuff, R. R. (1967). *Toward effective counseling and psychotherapy: Training and practice*. Hawthorne, NY US: Aldine Publishing Co.

Vader, A. M., Walters, S. T., Prabhu, G. C., Houck, J. M., & Field, C. A. (2010). The language of motivational interviewing and feedback: Counselor language, client language, and client drinking outcomes. *Psychology of Addictive Behaviors, 24*(2), 190-197.

Van Horn, D. H. A., & Bux, D. A. (2001). A pilot test of motivational interviewing groups for dually diagnosed inpatients. *Journal of Substance Abuse Treatment, 20*(2), 191-195.

Vasilaki, E. I., Hosier, S. G., & Cox, W. M. (2006). The efficacy of motivational interviewing as a brief intervention for excessive drinking: a meta-analytic review. *Alcohol and Alcoholism, 41*(3), 328-335.

Waldron, H. B., & Turner, C. W. (2008). Evidence-based psychosocial treatments for adolescent substance abuse. *Journal of Clinical Child and Adolescent Psychology, 37*(1)*,* 238-261.

Walker, D., Stephens, R., Rowland, J., & Roffman, R. (2011). The influence of client behavior during motivational interviewing on marijuana treatment outcome. *Addictive Behaviors, 36*(6), 669-673. doi: 10.1016/j.addbeh.2011.01.009

Walker, D. D., Roffman, R. A., Stephens, R. S., Berghuis, J. A., & Kim, W. (2006). Motivational enhancement therapy for adolescent marijuana users: A preliminary randomized controlled trial. *Journal of Consulting and Clinical Psychology, 74*(3), 628-632.

Walters, S. T. (2000). In praise of feedback: an effective intervention for college students who are heavy drinkers. *Journal of American College Health, 48*(5), 235-238.

Walters, S. T., Bennett, M. E., & Miller, J. H. (2000). Reducing alcohol use in college students: a controlled trial of two brief interventions. *Journal of Drug Education, 30*(3), 361-372.

Walters, S. T., Ogle, R., & Martin, J. E. (2002). Perils and possibilities of group-based motivational interviewing. In W. R. Miller & S. Rollnick (Eds.), *Motivational interviewing: Preparing people for change* (2nd ed., pp. 377–390). New York: Guilford Press.

Walters, S. T., Vader, A. M., Harris, T. R., Field, C. A., & Jouriles, E. N. (2009). Dismantling motivational interviewing and feedback for college drinkers: A randomized clinical trial. *Journal of Consulting and Clinical Psychology, 77*(1), 64-73.

Weiss, R., Jaffee, W., Menil, V., & Cogley, C. (2004). Group therapy for substance use disorders: What do we know? *Harvard Review of Psychiatry, 12*(6), 339-350.

Whaley, A. L. (1999). Preventing the high-risk sexual behavior of adolescents: focus on HIV/AIDS transmission, unintended pregnancy, or both? *The Journal of Adolescent Health: Official Publication of the Society for Adolescent Medicine, 24*(6), 376-382.

White, H. R., Morgan, T. J., Pugh, L. A., Celinska, K., Labouvie, E. W., & Pandina, R. J. (2006). Evaluating two brief substance-use interventions for mandated college students. *Journal of Studies on Alcohol, 67*(2), 309-317.

Yalom, I. (1985). *The Theory and Practice of Group Psychotherapy* (3rd ed.). New York: Basic Books.