Title

Do what you say and say what you are going to do: A preliminary meta-analysis of client change and sustain talk subtypes in motivational interviewing

Authors

Magill, Molly
Bernstein, Michael H
Hoadley, Ariel
et al.

Do what you say and say what you are going to do: A preliminary meta-analysis of client change and sustain talk subtypes in motivational interviewing

Molly Magill, Michael H. Bernstein, Ariel Hoadley, Brian Borsari, Timothy R. Apodaca, Jacques Gaume & J. Scott Tonigan

View supplementary material ⧉

Published online: 28 Jun 2018.

Submit your article to this journal ⧉

View Crossmark data ⧉

EMPIRICAL PAPER

Do what you say and say what you are going to do: A preliminary meta-analysis of client change and sustain talk subtypes in motivational interviewing

MOLLY MAGILL[1], MICHAEL H. BERNSTEIN[1], ARIEL HOADLEY [2],
BRIAN BORSARI[3], TIMOTHY R. APODACA[4], JACQUES GAUME[5], &
J. SCOTT TONIGAN [6]

[1]*Center for Alcohol and Addiction Studies, Brown University, Providence, RI, USA;* [2]*School of Public Health, Brown University, Providence, RI, USA;* [3]*San Francisco Veterans Affairs Health System and Department of Psychiatry, University of San Francisco, CA, USA;* [4]*Children's Mercy Kansas City, University of Missouri—Kansas City School of Medicine, Kansas City, MI, USA;* [5]*Lausanne University Hospital, Lausanne, Switzerland &* [6]*Center on Alcoholism, Substance Abuse, and Addictions, University of New Mexico, Albuquerque, NM, USA*

Abstract
Objective: This meta-analysis examines the predictive validity of client change language subtypes in motivational interviewing (MI) sessions addressing addictive behavior change. **Method:** A systematic review identified $k = 13$ primary studies, contributing 16 MI conditions ($N = 1556$). The pooled correlation coefficient was used to assess the significance, direction, and strength of seven language subtypes (i.e., reason, desire, need, ability, commitment, taking steps, and other) by three valences (i.e., frequency positive or *change talk*, frequency negative or *sustain talk*, and proportion change talk) and their relationship to subsequent engagement in addictive behavior. **Results:** For frequency measures, more sustain talk related to reason, desire, ability, and other were associated with more addictive behavior at follow up. Other change talk was associated with MI outcomes but in an unexpected direction (i.e., more addictive behavior). Proportion measures showed more proportion change talk-reason and -other statements were associated with less addictive behavior at follow up. Sensitivity analyses indicated some heterogeneity and instability of effect sizes, but no evidence of publication bias. **Conclusions:** This preliminary meta-analysis suggests that aggregate measures of change and sustain talk are comprised of statement subtypes that are not equally meaningful in predicting outcome following MI for addictive behavior change.

Keywords: motivational interviewing; change talk; sustain talk; technical hypothesis; meta-analysis

Clinical or methodological significance of this article: This preliminary meta-analysis suggests that aggregate measures of change and sustain talk are comprised of statement subtypes that are not equally meaningful in predicting outcome following motivational interviewing (MI) for addictive behavior change. Future MI research should regularly examine these language subtypes in process analyses. In addition, clinicians in MI are engaging in strategies that target motivational enhancement, but our results suggest that attendance to client ambivalence and, particularly, to client sustain talk are important to this process.

Introduction

In the field of MI, there has been a growing emphasis on what the client says within the interview as a prognostic factor, or even as a mechanism, of client outcome. Specifically, the MI technical model, or Technical Hypothesis (Miller & Rose, [2009]), posits that MI-consistent skills (e.g., open-ended questions, affirmations, simple and complex reflections)

influence how clients talk about behavior change. Subsequently, client statements in favor of behavior change (i.e., change talk) are theorized to be mechanisms of motivational enhancement (Amrhein, 2004; Miller & Rollnick, 2002; Miller & Rose, 2009). Two guiding frameworks for understanding MI process are self-determination and self-perception theories. Self-determination theory provides a context for conceptualizing the origins of human motivation, arguing that human beings are inherently driven toward healthy functioning and growth (Ryan & Deci, 2000). In self-perception theory, the way we know ourselves – our attitudes, values, and goals – is shaped by self-observation (Bem, 1972). In Miller and Rollnick's (2002) formulation, the MI client will *talk them self into changing* by hearing their own responses to content elicited, and selectively reinforced, by the interviewer. Finally and by translation, frequent statements against behavior change (i.e., sustain talk) will relate to worse outcomes after a motivational interview, and therefore, should not be reinforced during an MI session.

Over the past decade, the MI Technical Hypothesis has been the subject of numerous process research studies. A meta-analysis across 36 of these studies found support for some of the hypothesized technical pathways, and not others (Magill et al., 2018). Specifically, the review found significant effect sizes for four of five theorized paths from therapist technical skills to client change language and for two of three theorized paths from client change language to client outcome. In that study, client sustain talk or proportion of change talk (i.e., a combined measure of change and sustain talk) showed predictive validity with respect to follow-up risk behavior. Therefore, while the theory has prioritized the role of change talk, the empirical data mostly points to the importance of sustain talk, as well as the amount of client change talk relative to the total number of motivational statements, which could be a marker of client ambivalence (Magill et al., 2014; 2018). Statistical significance for sustain talk and non-significance for change talk was similarly found in a meta-analysis of 19 studies by Pace et al. (2017). However, there was heterogeneity in the pooled effect sizes reported for both of these meta-analyses, and when one looks at individual studies, the results are indeed mixed. Some studies show predictive/prognostic and mediation/mechanistic effects for change talk (Barnett et al., 2014; Houck, Manuel, & Moyers, 2018; Moyers, Martin, Houck, Christopher, & Tonigan, 2007; Pirlott, Kisbu-Sakarya, DeFrancesco, Elliot, & MacKinnon, 2012) and others point to the importance of sustain talk (Apodaca et al., 2014; Gaume et al., 2016a). Without a clearer understanding of the key process

variable/s in MI, we are poorly equipped to guide MI practitioners about optimal session foci and best practices for time management. When behavioral interventions are brief, lasting only a single session, these are particularly important implementation questions to address.

In prior reviews of the relationship between client language and client outcomes in MI, aggregate measures of change and sustain talk have been the focus of analyses. However, it is possible that certain underlying subtypes of these constructs (e.g., commitments to change versus reasons to change) have greater predictive validity than others. This speculation has borne out in some individual studies (Amrhein, Miller, Yahne, Palmer, & Fulcher, 2003; Baer et al., 2008; Gaume, Bertholet, Faouzi, Gmel, & Daeppen, 2010), but the relationship between client language subtypes and outcomes has never been examined in the context of meta-analysis. Therefore, the purpose of the present study was to test the relationship between seven subtypes of client language (i.e., reason, desire, need, ability, commitment, taking steps, other; see Supplemental Table 1 for definitions and examples) and client addictive behaviors at follow up. We examined not only language subtypes, but also language valence (i.e., change talk, sustain talk, and proportion change talk). Since the study is exploratory, no a priori hypotheses were established.

Methods

This study is a subsequent analysis of data compiled in a large-scale meta-analysis of process research in MI (AA023662; PI: Magill). The current analyses focus on 7 client language subtypes and 3 valence types to result in 21 types of effects across 13 primary studies that contributed 16 MI conditions. This report, therefore, focuses on the *b* path (mediator/mechanism to outcome) of the Technical Hypothesis of MI efficacy. Complete details regarding data acquisition have been reported elsewhere (Magill et al., 2018), and are briefly summarized here.

Primary Study Inclusion

Studies were eligible for inclusion if (a) they examined MI, Motivational Enhancement Therapy (MET), or Brief Motivational Intervention (BMI; including the Brief Alcohol Screening and Intervention for College Students [BASICS; Dimeff, 1999]) and (b) MI sessions were observationally coded (e.g., Houck, Moyers, Miller, Glynn, & Hallgren, 2010; Miller, Moyers, Ernst, & Amrhein, 2003) for

subsequent process analysis and (c) the targeted population was adolescents or adults, receiving intervention for behavior change and (d) the studies were published or in-press in a peer-reviewed, English-language journal between 2000 and 2017.

Literature Search

A literature search was conducted up to June of 2016, and was re-conducted in December to encompass up to January 1 of 2017. The search strategy included PsycINFO, PubMed, and Medline and used the following keywords: "change talk," "sustain talk," "client speech," "client language," "change language," "commitment language," "motivational interviewing skills," "motivational interviewing process," "motivational interviewing mediators," "motivational interviewing mechanisms," and "motivational interviewing ingredients." A hand search of included studies' reference lists and pertinent review papers (e.g., Apodaca & Longabaugh, 2009; Longabaugh, Magill, Morgenstern, & Huebner, 2013; Magill et al., 2014; Miller & Rollnick, 2013; Miller & Rose, 2009; Romano & Peters, 2015; Romano & Peters, 2016) was also conducted. The final step was a call for in-press papers to identified experts in MI process research. Figure 1 shows study inclusion, consistent with QUORUM guidelines (Moher et al., 1999). Eligibility was initially discussed between the first and third authors, but consensus was later provided by the research team.

Data Collection

For each eligible study, a trained research assistant compiled all relevant data. Effect sizes were taken from Pearson moment correlation matrices or computed from raw data that were requested from primary study authors. This approach allowed the maximum number of individual effect sizes as well as types of effect sizes (i.e., effect sizes based on frequencies vs proportions), regardless of what was reported in the primary study publication. Overall, 34 of 36 authors provided the requested data. Of these, 13 studies with 16 MI conditions (i.e., Apodaca, Magill, Longabaugh, Jackson, & Monti, 2013; Borsari et al., 2015; Davis, Houck, Rowell, Benson, & Smith, 2016; Vader, Walters, Prabhu, Houck, & Field, 2010 provided two MI conditions per study) were included in the present report. These eligible studies contained the required client language subtype data. Trained research assistants entered all correlation effect sizes into
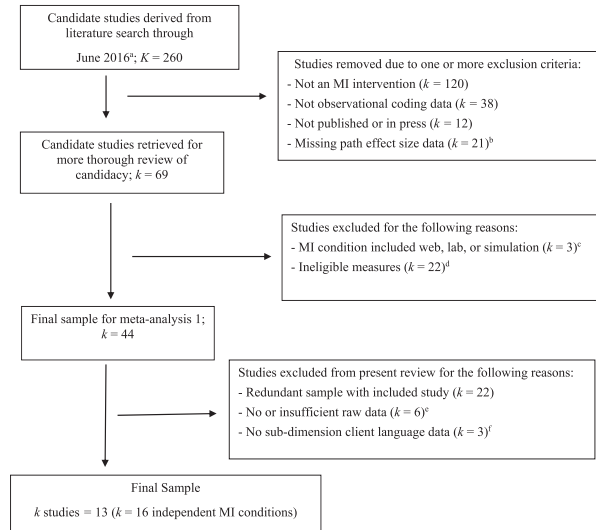

Figure 1. Flow of primary study inclusion.
Notes: K/k is defined as number of groups. (a) The June 2016 literature search was re-conducted to encompass up to 1 January 1 2017. However, no eligible studies, with independent samples, were derived at this time point. (b) E.g., Study examined MI skills in relation to outcomes. (c) E.g., Feldstein Ewing, Filbey, Sabbineni, Chandler, and Hutchison (2011); Glynn and Moyers (2010); Klonek, Lehmann-Willenbrock, and Kauffeld (2014). (d) E.g., Laws et al. (2015) [Physicians as Counselors Coding System]. (e) E.g., Amrhein et al. (2003); Morgenstern et al. (2012). (f) Campbell, Adamson, and Carter (2010); Kaplan, Keeley, Engel, Emsermann, and Brody (2013); and Pirlott et al. (2012).

Comprehensive Meta-Analysis spreadsheet software (Version 2), which was verified by a biostatistician staff member.

Analytic Plan

Overview. Descriptive analyses of primary study characteristics were conducted. Next, we examined effect sizes for the paths between client change language and client outcome at follow up. Finally, we assessed effect size stability, heterogeneity, and publication bias in sensitivity analyses.

Effect sizes of interest. We computed effect sizes for up to 21 different relationships within each study. Two key variations on language to outcome effects were of interest: (1) subtype (reason, desire, need, ability, commitment, taking steps, other) and (2) valence (i.e., change talk, sustain talk, proportion change talk). To contrast with language subtypes, aggregate measure pooled effect sizes for the current sample, were also reported.

Outcomes of interest. Target outcomes varied among primary studies, and multiple dependent

variables were reported by most studies. Therefore, we developed a ranking system for outcome variable selection. In order of preference, our primary outcomes were: (1) frequency (e.g., number of using days), (2) heavy frequency/quantity (e.g., number of heavy drinking days; defined as 4 or more drinks per day for women and 5 or more drinks per day for men; average amount consumed on a given using day), and (3) other outcomes (e.g., rates of improvement; consequences associated with use). For studies with multiple follow-ups, we prioritized: (1) early (i.e., 1–3 months), (2) mid-length (i.e., 4–6 months), and (3) late (i.e., 7+ months) time points. Finally, the majority of studies (11 of 13) reported outcomes in terms of addictive behavior frequency (e.g., number of using days), and on the few occasions when reported outcomes were positive (e.g., percentage of days abstinent), these effects were reverse scored.

Table I. Meta-analytic sample-level descriptive information.

Variable	Mean (SD)	Percent (k)
Client demographic factors		
Age	25.3(9.9)	
Percent female in sample	37.2(20.1)	
Sample type		
Adult sample		37.5(6)
College/young adult sample		43.8(7)
Adolescent sample		18.8(3)
Race/ethnicity		
Percent Caucasian in sample	68.8(24.0)	
Percent African American in sample	14.2(13.4)	
Percent Hispanic/Latino in sample	10.7(7.1)	
Client clinical factors		
Treatment-seeking sample		25.0(4)
Non-treatment-seeking sample		75.0(12)
Outcome of interest		
Alcohol study		75.0(12)
Other drug study		18.8(3)
Gambling study		6.3(1)
Treatment factors		
Session time in minutes	40.4(14.3)	
Treatment type		
Motivational Interviewing		18.8(13)
Motivational Enhancement Therapy		6.3(1)
Brief Alcohol Screening Intervention for College Students		12.5(2)
Setting type		
Specialty mental health/substance use		18.8(3)
College		31.3(5)
Medical		25.0(4)
Other		25.0(4)

Notes: Total k is 13 primary studies contributing 16 effect sizes. The coding of study characteristics was completed by two independent raters, with consensus review completed with the first author. All discrepancies were resolved in reference to the study report.

Effect size estimation and sensitivity analyses.
The pooled correlation coefficient effect size provides an inverse-variance-weighted indicator of the significance, strength, and direction of a bivariate relationship. Path effect sizes were modeled as random effects, where both known and unknown sources of effect size variability were assumed (Higgins & Thompson, 2002). Sensitivity analyses were conducted to examine the heterogeneity and stability of the effect sizes. To assess heterogeneity, the Q statistic was calculated and when the Q value was statistically significant, a measure of between-study variance was also reported (I^2). To assess stability, we looked for influential studies that, when removed, would change the statistical significance of the pooled effect estimate (i.e., a changed threshold of greater or less than .05; Baujat, Mahé, Pignon, & Hill, 2002). Where indicated, we provide a more conservative, trimmed estimate, with influential studies removed. Finally, when pooled effect sizes were significant, a publication bias analysis was conducted (Rosenthal, 1979). In summary, the resulting pooled correlation coefficient effect sizes indicate the significance, strength, and magnitude of the relationship between the 21 language subtypes and addictive behavior at follow-up post MI (i.e., effect magnitude benchmarks can be interpreted as follows: $r \leq .2$ is "small"; $r \leq .4$ is "medium"; and $r \geq .5$ is "large," Cohen, 1988).

Results

Primary Study Sample

Study-level descriptive data are provided in Supplemental Table 2 and meta-analytic sample-level descriptive data are provided in Table I. A total of 13 studies treated $N = 1556$ individuals. On average, study samples had 97 participants ($SD = 63$; median = 91). Most often the studies treated young adults ($k = 7$), followed by adults ($k = 6$), and adolescents ($k = 3$; overall M age = 25, $SD = 9.9$). Studies had primarily male (63%) and Caucasian (69%) samples and racial or ethnic representation was 14% African American and 11% Hispanic/Latino, on average. These studies targeted mostly non-treatment seeking (75%), alcohol or other drug users (one gambling study; Hodgins, Ching, & McEwen, 2009). The MI, BASICS, and MET sessions were typically 40 minutes in length ($SD = 14$ minutes). The MI interventions were primarily manualized, and there was a fairly equal distribution across treatment setting types (i.e., specialty mental health/substance use, college, medical, and other). Finally, across this sample of studies, global empathy and MI Spirit were "good" on average

($M = 4.6$ [$SD = 0.5$] and $M = 4.3$ [$SD = 0.6$], respectively) on a 5-point scale.[1]

The Average Effect of Client Language on Outcome—By Subtype and Valence

The significance, magnitude, and direction of pooled effect sizes varied by subtype and valence in this meta-analytic review. For frequency of change talk statements, the overall random effects, pooled effect size was non-significant and homogeneous ($r = .033$, $k = 16$: 95% CI [−.035, .100]; $p = .339$; $Q > .05$), as were frequency of reason, desire, need, commitment, and taking steps statements. Change talk-ability statements were non-significant but were heterogeneous ($r = −.037$; 95% CI [−.163, .090]; $p = .566$; $Q < .05$, $I^2 = 79\%$). For change talk-other statements, the effect was small, positive, significant, and homogeneous ($r = .074$; 95% CI [.006, .142]; $p = .034$; $Q > .05$). In other words, these types of statements were related to more follow-up addictive behavior in MI. This effect became non-significant if one of several influential studies was removed (trimmed r's .045 to .075, p's > .05; Apodaca et al., 2014; Apodaca et al., 2014 [both conditions]; Borsari et al., 2015 [site 1]; Gaume et al., 2016a; Vader et al., 2010 [MI condition only]).

For frequency of sustain talk statements, the overall random effects, pooled effect size was significantly related to the greater frequency of addictive behavior and homogenous ($r = .203$, $k = 16$: 95% CI [.151, .253]; $p < .001$; $Q > .05$). Subtype results for sustain talk also showed positive, significant, and homogeneous effect sizes for four of seven indicators. Specifically, more sustain talk statements related to reasons, ability, desire, and other were associated with more addictive behavior at follow up, while need, taking steps, and commitment statements were non-significant. Among these latter subtypes, sustain talk-need statements were moderately heterogeneous ($Q < .05$; $I^2 = 49\%$). Significant sustain talk effect sizes were generally small in magnitude ($r = .062$ for sustain talk-ability to $r = .155$ for sustain talk-reasons). Finally, in the analyses of effect size stability, two influential studies resulted in marginally non-significant effects for sustain talk-ability (trimmed $r = .055$, $p = .063$; Apodaca et al., 2014 [MI condition]; trimmed $r = .056$, $p = .055$; Gaume et al., 2010).

For proportion of total motivational statements that were change talk, the pooled effect size was significantly related to less follow-up addictive behavior ($r = −.167$, $k = 15$: 95% CI [−.241, −.091]; $p < .001$; $Q < .05$, $I^2 = 44\%$). Subtype results showed higher proportion change talk – reason statements were

significantly related to less addictive behavior ($r = −.089$; 95% CI [−.167, −.009]; $p = .029$), and this effect was small and moderately heterogeneous ($Q > .05$, $I^2 = 47\%$). For proportion change talk – reason statements, the effect became non-significant if one of three influential studies was removed (trimmed $r = −.073$, $p = .081$; Apodaca et al., 2014; trimmed $r = −.085$, $p = .053$; Gaume, Gmel, & Daeppen, 2008; Gaume, Gmel, Faouzi, & Daeppen, 2008; Gaume, Gmel, Faouzi, & Daeppen, 2009; trimmed $r = −.081$, $p = .052$; Hodgins et al., 2009). Finally, proportion change talk-other statements were related to less addictive behavior ($r = −.109$; 95% CI [−.177, −.040]; $p = .002$), and this effect was small, homogenous, and showed no influential studies. The remaining subtypes were not significant. For a summary view of pooled effect sizes by subtype, please see Figure 2, which illustrates 21 effects organized by valence.[2]

Discussion

The present meta-analysis examined client motivational statements within MI sessions and their relationship to addictive behavior at follow up. The predictive validity of language subtypes, rather than aggregate measures of change and sustain talk, was the primary focus of this report. Overall, 6 of 21 correlation paths were supported in the direction suggested by the MI technical model (Miller & Rose, 2009), and one path had a significant effect in a direction opposite of what might be expected. Specifically, of change talk subtypes, only other statements were statistically significant and were surprisingly related to worse, rather than better MI outcome. Worse MI outcomes were also associated with reason, desire, ability, and other sustain talk subtypes. Finally, the proportion of change talk-reason and -other statements were associated with better MI outcomes at follow up, suggesting that when positive behavioral intentions are of interest, these combined change and sustain talk indicators might have optimal predictive validity.

Change Talk Has Unreliable Predictive Validity—What Clients Say They Will Do

Among change talk subtypes, only other statements predicted MI outcomes. However, these statements were associated with greater addictive behavior frequency, which is contrary to what MI theory might suggest. The remaining subtypes (i.e., reason, desire, need, ability, commitment, taking steps) were not significantly related to outcomes in MI. Our results show poor overall predictive validity for
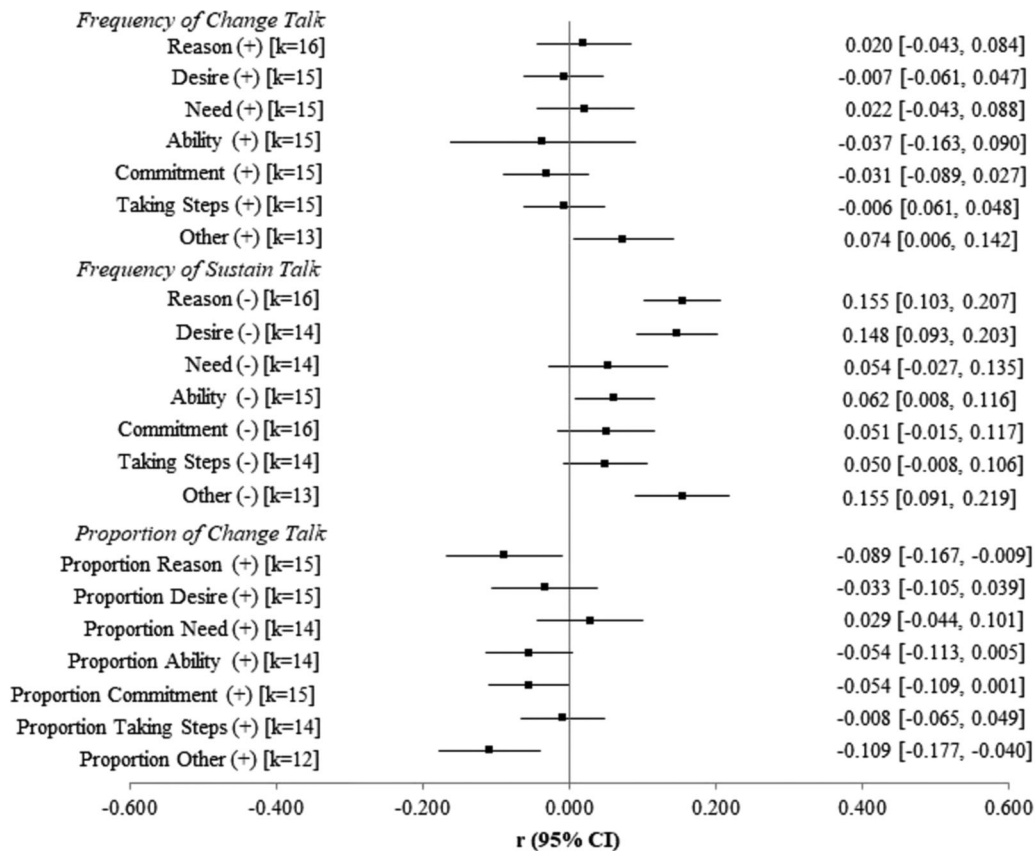
Figure 2. Pooled correlation coefficients and 95% confidence intervals for language subtypes and risk behavior outcome.

client statements in favor of change, regardless of the subtype examined. Why might this be? Perhaps some clients are more genuine when talking about changing addictive behavior in an MI session, while other clients are disingenuous or particularly ambivalent. In this meta-analysis, the majority of participants were non-treatment seeking, older adolescent or young adult alcohol users. Therefore, it may be that change talk predictive validity varies by client factors. Miller and Rollnick (2013) have another take on this question. The authors speak to a relational phenomenon where clients could be "telling you what [*the client thinks*] you want to hear," and they refer to this as *dubious change talk*. Miller and Rollnick (2002) also view dubious change talk as sometimes "wishful thinking." In either case, the authors suggest the clinician should seek depth in these statements via continued exploration, and the current results suggest this is particularly true for other statements.

Unreliable predictive validity for change talk subtypes might also relate to a measurement factor – the use of a language frequency metric (i.e., how much a given statement occurs), rather than language strength (i.e., the intensity or conviction in the statement). While consistent with the notion that MI

therapists should help the client talking them self into changing, maybe how much isn't the best indicator. For example, consider the following two statements: (1) "I guess my drinking is creating problems for me" (low strength, change talk-reason), versus (2) "I hate what drinking has done to me!" (high strength, change talk-reason), which have the same frequency value, but differ in strength. Over 15 years ago, MI researchers observed that it was not the occurrence of verbal commitment, but its strength, that was prognostic of change (Amrhein et al., 2003). Strength measures are rarely utilized in MI process studies because these data are not collected and/or have poor reliability (e.g., Houck et al., 2010). That said, Gaume et al. (2016b) demonstrated that strength, and not frequency, in an MI session was associated with improved drinking outcomes among non-treatment seeking young men. MI process research should consider if strength rating methodology can be improved. However, if too fraught with subjectivity, then computer-based (Atkins, Steyvers, Imel, & Smyth, 2014) or neuroscientific (Houck, Moyers, & Tesche, 2013) methods might be the pivotal next steps toward reliably capturing a client's true positive intention regarding behavior change.

Sustain Talk Has Consistent Predictive Validity—What Clients Say They Won't Do

Among sustain talk subtypes, reasons, abilities, desires, and other predicted outcome in this meta-analysis, while taking steps and commitment did not, which is counter-intuitive to what one might expect clinically. In other words, sustain talk statements, related to what Amrhein (2004) refers to as "preparatory" thinking, were associated with less change at follow up while sustain talk "action" statements were not. Negative commitments or reported steps can be quite rare in MI process studies (e.g., Apodaca et al., 2014; Kahler et al., 2016; Vader et al., 2010). As such, our results could relate to low power to detect these statements' effects. To say, "I won't stop smoking weed" (sustain talk-commitment) or "I drank last week" (sustain talk-taking steps) requires a high level of confidence and/or trust on the part of the speaker, and that might be an unrealistic expectation for some clients in a single-session, brief intervention. In general, sustain talk occurs less frequently than change talk; it is *pulled for* (i.e., evoked by the clinician) only in certain MI session components (i.e., Pros and Cons, and to a lesser extent Change Plan) and it is very likely, less prone to social desirability bias.

It might be useful to incorporate sustain talk subtypes and timing when considering the validity of client negative intention in MI. For example, a recent study found that sustain talk occurred at a greater frequency during the earlier portion of MI sessions among those who reported greater substance use outcomes at follow up compared to those who did not use substances (Rodriguez, Walters, Houck, Ortiz, & Taxman, 2018). In this meta-analysis, we examined the frequency of sustain talk and its subtypes, but we did not prioritize one portion of the session over another. A few process studies have highlighted the value of making a temporal distinction between client language that arises during *engaging and focusing* and talk that arises during MI *planning* (Amrhein et al., 2003; Bertholet, Faouzi, Gmel, Gaume, & Daeppen, 2010; Kahler et al., 2016; Morgenstern et al., 2017). A question to consider in the future is whether greater sustain talk represents a prognostic patient factor (e.g., a defensive or reactant trait) occurring early in the session, a resistance marker that emerges later via the MI process (i.e., a relational state), or both? In other words, it is possible that the total frequency of sustain talk-reason, desire, ability, and other statements are less relevant than the timing of when they occur. If future studies more regularly made this distinction, the field could become better equipped to guide clinicians in timing their work with sustain talk in MI.

The Other Talk Category—What Is It?

Results on the predictive role of the other subtype across valence categories are challenging to interpret because of the intended purpose of this category within MI coding measures and procedures. In the present meta-analysis, other sustain talk was related to worse outcome, proportion other change talk was related to better outcome, and other change talk was associated with outcome in an unexpected direction (i.e., related to poorer outcomes, albeit weakly). The other language subtype is designed to capture hypothetical language, indirect statements, broad attitudes, and other statements that do not sufficiently qualify as another subtype (Houck et al., 2010; Miller et al., 2003; 2008a). These statements are also often depersonalized or non-committal, such as, "If I could just be on a desert island for a month, I could quit" (change talk-other; Miller, Moyers, Manuel, Christopher, & Amrhein, 2008b, p. 18), and by nature *should capture ambivalence*. As such, these are the statements that should have relatively minimal value in predicting change. Given the results found here, future research should attempt to define this category further, since there may be some types of statements embedded within this grouping that have value in representing individuals' intentions. Moreover, because our sample was primarily adolescents and young adults not seeking treatment for behavior change, it could be that these remarks are more common among younger clients. Therefore, how other talk functions across developmental stages should be considered in future MI process research.

Limitations

This meta-analysis has some key limitations to consider. Most importantly, the size of the primary study sample is relatively small and the majority of studies targeted young, non-clinical samples with problematic use rather than an addictive disorder. That said, these are typical samples for brief intervention with MI. Next and given some pooled effect sizes showed heterogeneity or influential studies, the results reported here should be considered preliminary and should be confirmed with a larger sample. Another limitation is possibly poor or fair interrater reliability of certain language subtypes within pooled estimates. Among the studies in this report, we were able to obtain interrater reliability data for 80%, and these Intraclass Correlation Coefficients (ICC's) ranged from "poor" to "excellent."[3] Typically, poorly rated items are infrequently occurring items (Hallgren, 2012), as was the case in this meta-analysis. The use of the correlation coefficient, which assumes a

linear and bivariate relationship between client language and the targeted outcome, may be another limitation. Finally, three subtype effect sizes were statistically heterogeneous and these were: change talk-ability, sustain talk-need, and proportion change talk-reason. While this suggests that the majority of pooled effect sizes did not contain systematic variability (Higgins, Thompson, Deeks, & Altman, 2003), it hindered our ability to consider moderating or sub-group variables of potential substantive interest. Some candidate indictors would be researcher allegiance, sample representativeness, and client preferences/treatment-seeking status (Spielmans & Flückiger, 2018). With the noted limitations in mind, the following conclusions can be considered.

Conclusions

This preliminary meta-analysis suggests that aggregate measures of change and sustain talk are comprised of statement subtypes that are not equally meaningful in predicting subsequent follow-up behavior. As a result, greater attention to, and thus better reporting of, change statement subtypes is important for future MI process research. Moreover, sustain talk shows more consistent predictive validity than change talk. Therefore, while change talk has received the most attention in the clinical literature, this study suggests that it only has prognostic value regarding client intentions for addictive behavior change when combined with sustain talk.

Acknowledgement

Funding

Notes

1 Global scores on a 7-point scale from early versions of the Motivational Interviewing Skill Code (Miller et al., 2003, 2008a) were transformed to a 5-point scale. Therefore, ratings of "average" or lower and "good" or higher were consistent across studies.

2 Analyses of possible publication bias showed that over 50 unpublished null studies would be required to raise the observed, significant, *p* values to above alpha of .05. Given the size of this number relative to published studies, we suggest no evidence of publication bias.

3 Percent of studies reporting change talk subtype with poor ICC value (Shrout & Fleiss, 1979): reason – 0%; ability – 18%; need – 22%; desire – 27%; commitment – 9%; taking steps – 11%; other – 0%. Percent of studies reporting sustain talk subtype with poor ICC value (Shrout & Fleiss, 1979): reason – 0%; ability – 27%; need – 22%; desire – 20%; commitment – 0%; taking steps – 14%; other – 44%.

ORCID

Ariel Hoadley ⓘ http://orcid.org/0000-0003-1360-0358
J. Scott Tonigan ⓘ http://orcid.org/0000-0002-6668-2038

Supplemental data

References

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

Amrhein, P. C. (2004). How does motivational interviewing work? What client talk reveals. *Journal of Cognitive Psychotherapy*, *18*(4), 323–336. doi:10.1891/jcop.18.4.323.64001

Amrhein, P. C., Miller, W. R., Yahne, C., 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. doi:10.1037/0022-006X.71.5.862

Apodaca, T., & Longabaugh, R. (2009). Mechanisms of change in motivational interviewing: A review and preliminary evaluation of the evidence. *Addiction*, *104*(5), 705–715. doi:10.1111/j.1360-0443.2009.02527.x

*Apodaca, T. R., Borsari, B., Jackson, K. M., Magill, M., Longabaugh, R., Mastroleo, N. R., & Barnett, N. P. (2014). Sustain talk predicts poorer outcomes among mandated college student drinkers receiving a brief motivational intervention. *Psychology of Addictive Behaviors*, *28*(3), 631–638. doi:10.1037/a0037296

*Apodaca, T. R., Magill, M., Longabaugh, R., Jackson, K. M., & Monti, P. M. (2013). Effect of a significant other on client change talk in motivational interviewing. *Journal of Consulting and Clinical Psychology*, *81*(1), 35–46. doi:10.1037/a0030881

Atkins, D. C., Steyvers, M., Imel, Z. E., & Smyth, P. (2014). Scaling up the evaluation of psychotherapy: Evaluating motivational interviewing fidelity via statistical text classification. *Implementation Science*, *9*(1), 265. doi:10.1186/1748-5908-9-49

*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, 570–575. doi:10.1037/a0013022

Barnett, E., Moyers, T. B., Sussman, S., Smith, C., Rohrbach, L. A., Sun, P., & Spruijt-Metz, D. (2014). From counselor skill to decreased marijuana use: Does change talk matter? *Journal of Substance Abuse Treatment*, 46(4), 498–505. doi:10.1016/j.jsat.2013.11.004

Baujat, B., Mahé, C., Pignon, J. P., & Hill, C. (2002). A graphical method for exploring heterogeneity in meta-analysis: Application to a meta-analysis of 65 trials. *Statistics in Medicine*, 21(18), 2641–2652. doi:10.1002/sim.1221

Bem, D. J. (1972). Self-perception theory. *Advances in Experimental Social Psychology*, 6, 1–62. doi:10.1016/S0065-2601(08)60024-6

Bertholet, N., Faouzi, M., Gmel, G., Gaume, J., & Daeppen, J. B. (2010). Change talk sequence during brief motivational intervention, towards or away from drinking. *Addiction*, 105(12), 2106–2112. doi:10.1111/j.1360-0443.2010.03081.x

*Borsari, B., Apodaca, T. R., Jackson, K. M., Mastroleo, N. R., Magill, M., Barnett, N. P., & Carey, K. B. (2015). In-session processes of brief motivational interventions in two trials with mandated college students. *Journal of Consulting and Clinical Psychology*, 83(1), 56–67. doi:10.1037/a0037635

Campbell, S. D., Adamson, S. J., & Carter, J. D. (2010). Client language during motivational enhancement therapy and alcohol use outcome. *Behavioral and Cognitive Psychotherapy*, 38, 399–415. doi:10.1017/S1352465810000263

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

*Davis, J. P., Houck, J. M., Rowell, L. N., Benson, J. G., & Smith, D. C. (2016). Brief motivational interviewing and normative feedback for adolescents: Change language and alcohol use outcomes. *Journal of Substance Abuse Treatment*, 65, 66–73. doi:10.1016/j.jsat.2015.10.004

Dimeff, L. A. (ed.). (1999). *Brief alcohol screening and intervention for college students (BASICS): A harm reduction approach*. New York, NY: Guilford Press.

Feldstein Ewing, S. W., Filbey, F. M., Sabbineni, A., Chandler, L. D., & Hutchison, K. E. (2011). How psychosocial alcohol interventions work: A preliminary look at what fMRI can tell us. *Alcoholism: Clinical and Experimental Research*, 35(4), 643–651. doi:10.1111/j.1530-0277.2010.01382.x

Gaume, J., Magill, M., Mastroleo, N. R., Longabaugh, R., Bertholet, N., Gmel, G., & Daeppen, J. B. (2016b). Change talk during brief motivational intervention with young adult males: Strength matters. *Journal of Substance Abuse Treatment*, 65, 58–65. doi:10.1016/j.jsat.2016.01.005

*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. doi:10.1016/j.jsat.2010.06.010

*Gaume, J., Bertholet, N., Faouzi, M., Gmel, G., & Daeppen, J. B. (2013). Does change talk during brief motivational interventions with young men predict change in alcohol use? *Journal of Substance Abuse Treatment*, 44(2), 177–185. doi:10.1016/j.jsat.2012.04.005

*Gaume, J., Gmel, G., & Daeppen, J. (2008). Brief alcohol interventions: Do counsellors' and patients' communication characteristics predict change? *Alcohol & Alcoholism*, 43(1), 62–69. doi:10.1093/alcalc/agm141

*Gaume, J., Gmel, G., Faouzi, M., & Daeppen, J. (2008). Counsellor behaviours and patient language during brief motivational interventions: A sequential analysis of speech. *Addiction*, 103(11), 1793–1800. doi:10.1111/j.1360-0443.2008.02337.x

*Gaume, J., Gmel, G., Faouzi, M., & Daeppen, J. (2009). Counselor skill influences outcomes of brief motivational interventions. *Journal of Substance Abuse Treatment*, 37(2), 151–159. doi:10.1016/j.jsat.2008.12.001

*Gaume, J., Longabaugh, R., Magill, M., Bertholet, N., Gmel, G., & Daeppen, J. B. (2016a). Under what conditions? Therapist and client characteristics moderate the role of change talk in brief motivational intervention. *Journal of Consulting and Clinical Psychology*, 84(3), 211–220. doi:10.1037/a0039918

Glynn, L. H., & Moyers, T. B. (2010). Chasing change talk: The clinician's role in evoking client language about change. *Journal of Substance Abuse Treatment*, 39(1), 65–70. doi:10.1016/j.jsat.2010.03.012.

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. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3402032/

Higgins, J., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, 21(11), 1539–1558. doi:10.1002/sim.1186

Higgins, J., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses [journal article as teaching resource, deposited by John Flynn]. *British Medical Journal*, 327, 557–560. doi:10.1136/bmj.327.7414.557

*Hodgins, D. C., Ching, L. E., & McEwen, J. (2009). Strength of commitment language in motivational interviewing and gambling outcomes. *Psychology of Addictive Behaviors*, 23(1), 122–130. doi:10.1037/a0013010

Houck, J. M., Manuel, J. K., & Moyers, T. B. (2018). Short- and long-term effects of within-session client speech on drinking in the COMBINE study. *Journal of Studies on Alcohol and Drugs*, 79(2), 217–222.

Houck, J. M., Moyers, T. B., Miller, W. R., Glynn, L. H., & Hallgren, K. A. (2010). Motivational interviewing skill code (MISC) version 2.5.

Houck, J. M., Moyers, T. B., & Tesche, C. D. (2013). Through a glass darkly: Some insights on change talk via magnetoencephalography. *Psychology of Addictive Behaviors*, 27(2), 489–500. doi:10.1037/a0029896

*Kahler, C. W., Caswell, A. J., Laws, M. B., Walthers, J., Magill, M., Mastroleo, N. R., … Monti, P. M. (2016). Using topic coding to understand the nature of change language in a motivational intervention to reduce alcohol and sex risk behaviors in emergency department patients. *Patient Education and Counseling*, 99(10), 1595–1602. doi:10.1016/j.pec.2016.05.003

Kaplan, J. E., Keeley, R. D., Engel, M., Emsermann, C., & Brody, D. (2013). Aspects of patient and clinician language predict adherence to antidepressant medication. *The Journal of the American Board of Family Medicine*, 26(4), 409–420. doi:10.3122/jabfm.2013.04.120201

Klonek, F. E., Lehmann-Willenbrock, N., & Kauffeld, S. (2014). Dynamics of resistance to change: A sequential analysis of change agents in action. *Journal of Change Management*, 14(3), 334–360. doi:10.1080/14697017.2014.896392

Laws, M. B., Rose, G. S., Beach, M. C., Lee, Y., Rogers, W. S., Velasco, A. B., & Wilson, I. B. (2015). Patient-provider concordance with behavioral change goals drives measures of motivational interviewing consistency. *Patient Education and Counseling*, 98(6), 728–733. doi:10.1016/j.pec.2015.02.014

Longabaugh, R., Magill, M., Morgenstern, J., & Huebner, R. (2013). Mechanisms of behavior change in treatment for alcohol and other drug use disorders. In B.S. McCrady, & E.E. Epstein (Eds.), *Addictions: A comprehensive guidebook* (2nd ed, pp. 572–596). New York, NY: Oxford University Press.

Magill, M., Apodaca, T. R., Borsari, B., Gaume, J., Hoadley, A., Gordon, R., … Moyers, T. (2018). A meta-analysis of motivational interviewing process: Technical, relational, and conditional process models of change. *Journal of Consulting and Clinical Psychology*, 86(2), 140–157. doi:10.1037/ccp0000250

Magill, M., Gaume, J., Apodaca, T. R., Walthers, J., Mastroleo, N. R., Borsari, B., & Longabaugh, R. (2014). The technical hypothesis of motivational interviewing: A meta-analysis of MI's key causal model. *Journal of Consulting and Clinical Psychology*, 82(6), 973–983. doi:10.1037/a0036833

Miller, W. R., Moyers, T. B., Ernst, D., & Amrhein, P. (2003). *Manual for the motivational interviewing skill code version 2.0 (MISC)* (Unpublished Manual).

Miller, W. R., Moyers, T. B., Ernst, D., & Amrhein, P. (2008a). *Manual for the motivational interviewing skill code version 2.1 (MISC)* (Unpublished Manual).

Miller, W. R., Moyers, T. B., Manuel, J. K., Christopher, P., & Amrhein, P. (2008b). *Revision for client language coding: MISC 2.1: Client language assessment in motivational interviewing (CLAMI) Segment* (Unpublished Manual).

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

Miller, W. R., & Rollnick, S. (2013). *Motivational interviewing: Helping people change* (3rd ed.). New York, NY: Guilford Press.

Miller, W. R., & Rose, G. S. (2009). Toward a theory of motivational interviewing. *American Psychologist*, 64, 527–537. doi:10.1037/a0016830.

Moher, D., Cook, D. J., Eastwood, S., Olkin, I., Rennie, D., & Stroup, D. F. (1999). Improving the quality of reports of meta-analyses of randomized controlled trials: The QUOROM statement. *The Lancet*, 354, 1896–1900. doi:10.1046/j.1365-2168.2000.01610.x.

Morgenstern, J., Kuerbis, A., Amrhein, P., Hail, L., Lynch, K., & McKay, J. R. (2012). Motivational interviewing: A pilot test of active ingredients and mechanisms of change. *Psychology of Addictive Behaviors*, 26(4), 859–869. doi:10.1037/a0029674

Morgenstern, J., Kuerbis, A., Houser, J., Levak, S., Amrhein, P., Shao, S., & McKay, J. R. (2017). Dismantling motivational interviewing: Effects on initiation of behavior change among problem drinkers seeking treatment. *Psychology of Addictive Behaviors*, 31(7), 751–762. doi:10.1037/adb0000317

Moyers, T. B., Martin, T., Manuel, J. K., Miller, W. R., & Ernst, D. (2007). *The Motivational Interviewing Treatment Integrity (MITI) code: Version 3.0* (Unpublished manuscript). Albuquerque, NM: University of New Mexico.

*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. doi:10.1037/a0017189

Pace, B. T., Dembe, A., Soma, C. S., Baldwin, S. A., Atkins, D. C., & Imel, Z. E. (2017). A multivariate meta-analysis of motivational interviewing process and outcome. *Psychology of Addictive Behaviors*, 31(5), 524–533. doi:10.1037/adb0000280

Pirlott, A. G., Kisbu-Sakarya, Y., DeFrancesco, C. A., Elliot, D. L., & MacKinnon, D. P. (2012). Mechanisms of motivational interviewing in health promotion: A Bayesian mediation analysis. *International Journal of Behavioral Nutrition and Physical Activity*, 9(1), 69. doi:10.1186/1479-5868-9-69

Rodriguez, M., Walters, S. T., Houck, J. M., Ortiz, J. A., & Taxman, F. S. (2018). The language of change among criminal justice clients: Counselor language, client language, and client substance use outcomes. *Journal of Clinical Psychology*, 74, 626–636. doi:10.1002/jclp.22534

Romano, M., & Peters, L. (2015). Evaluating the mechanisms of change in motivational interviewing in the treatment of mental health problems: A review and meta-analysis. *Clinical Psychology Review*, 38, 1–12. doi:10.1016/j.cpr.2015.02.008

Romano, M., & Peters, L. (2016). Understanding the process of motivational interviewing: A review of the relational and technical hypotheses. *Psychotherapy Research*, 26(2), 220–240. doi:10.1080/10503307.2014.954154

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, 86(3), 638–641. doi:10.1037/0033-2909.86.3.638

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68–78. doi:10.1037/0003-066X.55.1.68

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86(2), 420–428. doi:10.1037/0033-2909.86.2.420

Spielmans, G. I., & Flückiger, C. (2018). Moderators in psychotherapy meta-analysis. *Psychotherapy Research*, 1–14. doi:10.1080/10503307.2017.1422214

*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. doi:10.1037/a0018749