Challenges in testing fidelity in Motivational Interviewing oral health interventions

T Newton[1] and L M Jamieson[2]

[1]Unit of Social and Behavioural Sciences, King's College London Dental Institute
[2]Australian Research Centre for Population Oral Health, School of Dentistry, The University of Adelaide, Australia

Correspondence to: Professor J T Newton, Unit of Social and Behavioural Sciences, King's College London Dental Institute, Bessemer Road, London, SE5 9RW.
Email: tim.newton@kcl.ac.uk

Objective: Motivational Interviewing (MI), a counselling technique designed to stimulate a client's inner will to change, is based on four principles: empathy, developing discrepancy between current and an alternate behaviour, reinforcing self-efficacy and rolling with resistance to change. The most important component of MI is thought to be the MI 'spirit' which is concerned with enhancing client collaboration as well a client's autonomy. Despite evidence of efficacy and increasing use of MI in a range of oral health settings, little attention has been paid to documenting the integrity of MI delivery (fidelity). This paper aims to explore the challenges associated with the assessment of MI fidelity. **Methods:** A summary of challenges associated with MI fidelity assessment in health intervention contexts will be provided; encompassing topics such as budgetary constraints, time constraints and expertise constraints. The importance of becoming familiar with the tools available to assess MI fidelity will be discussed as well as the various consortiums available through which to communicate with experts in the MI fidelity field. **Results:** The challenges to conducting appropriate fidelity assessment for oral health interventions involving MI are many and varied. However, if evidence regarding the utility of MI interventions in oral health practice is to be translated to policy, the clinical setting, teaching and research, there is a responsibility among all those involved in MI oral health interventions to responsibly consider fidelity assessment as a critical part in the research process. **Conclusions:** There is an opportunity for communication among oral health researchers with an interest in MI to explore pathways through which barriers in fidelity assessment in MI-based interventions may be overcome.

Key words: Motivational Interviewing, oral health, oral health interventions

This paper has three main purposes. The first is to reiterate the importance of assessing fidelity for motivational interviewing (MI) interventions (or for any other interventions, be they psychological or pharmacological etc). The second purpose is to describe particular methods that have been suggested for quantifying MI fidelity, while the final purpose is to address the main challenges to determining the fidelity of an intervention.

WHY FIDELITY?

Carroll and colleagues (2002) identified three main reasons why the testing of the fidelity of an intervention is important. First, and perhaps the most obvious, is the importance of assessing the integrity of the independent variable. In a trial of an interventional drug, we might assess such integrity by determining the level of circulating drug in the individual's bloodstream. So in testing the effect of a psychological intervention such as MI, we need to know whether participants have been exposed to the correct intervention at an appropriate level. This is inherently more difficult than it might first appear. Without being able to assess the degree to which a person displays the behaviours and skills necessary to engage in MI, we cannot determine whether training in such technique has been effective, or whether our trainees have reached satisfactory levels of performance. Being able to assess competence also enables us to ensure that the translation of techniques away from highly structured research settings to care settings is able to occur. It provides an answer to the fundamental question '*Is the loss of effectiveness that so often occurs when moving from controlled research settings to community based interventions the result of a loss of fidelity or the result of the broader determinants of behaviour and health?*'

As described, there are a number of tools available for assessing the fidelity of MI. Each differs in format and the types of measures that are assessed. There is additional variation in the intended use of each scale. A brief overview of the design and property of the most commonly used measures is provided:

1. Yale Adherence and Competency Scale (YACS) (Carroll *et al.*, 2000)

The YACS is a general system for rating practitioner adherence and competence in delivering behavioural treatments for substance use disorders. The system includes three scales measuring 'general' aspects of drug abuse treatment (assessment, general support, goals of treatment), as well as three scales measuring critical elements of three treatments that are frequently implemented as control or comparison treatments in clinical research in addiction (clinical management, twelve step facilitation and cognitive behavioural therapy). Validation of the YACS using data from a randomised clinical trial indicated that the scales have excellent reliability, factor structure, concurrent and discriminant validity. Correlations between adherence and competence scores within scales were in the moderate range, indicating independence (and thus non-redundancy) of

these dimensions.

It is important to note that the YACS was devised to assess the validity of a range of psychological interventions. It assesses both adherence and quality, is reasonably reliable across raters with training, and correlates well with other measures (for example, MISTS; Corvino *et al*., 2000). It is primarily intended as a research instrument.

2. Motivational Interviewing Skill Code (MISC) (Miller and Mount, 2001)

Miller and Mount (2001) developed the MISC in 2001 in a bid to assess specific domains of counsellor and client functioning within MI sessions. The MISC utilises three separate techniques for reviewing therapist competence in the use of MI, each gathered in a separate review or 'pass' of the session tape. First, global assessments are made of MI-relevant therapist and client characteristics using a seven-point Likert scale. For the therapist, six global characteristics are measured: acceptance, egalitarianism, empathy, genuineness, warmth and overall MI spirit. For the client, four global characteristics are measured: affect, cooperation, disclosure and engagement. Two characteristics of the interaction between the therapist and client are also assessed with global scores: benefit and collaboration.

In a second coding pass, specific behaviours are counted during MI sessions. For therapists, 27 behaviours are coded, including both those specific to MI (asking permission before giving advice) and those common to many different types of therapy (asking questions, reflections). Four types of client speech are counted, reflecting the importance of client language in MI sessions. Frequency counts of client speech about the possibility of changing (change talk) as well as resistance to change (resist talk) are made, as well as occasions where the client simply follows the therapist's requests for information (follow/neutral) or asks questions of their own. The third pass in the MISC measures the relative amount of time spent talking during the session by both the client and the therapist. The most common use of the MISC has been to document changes in therapist competence before and after training in MI (Moyers *et al*., 2005).

Originally designed for use in research, the MISC is poor on reliability and intensive to use, but appears to discriminate between experienced and less experienced MI therapists.

3. The Motivational Interviewing Process Code (MIPC) (Barsky and Coleman, 2001)

The MIPC has not been widely used and there are few data on its psychometric properties. The instrument comprises 25 items in two sub-scales; one assessing functional skills and the other assessing dysfunctional skills. In regards to reliability, rater agreement has been documented as 51% for functional and 75% for dysfunctional skills respectively. Construct validity has been assessed through a modified Delphi technique (Barsky and Coleman, 2001).

4. The Motivational Interviewing Treatment Integrity Scale (MITI) (Moyers *et al*., 2005)

The MITI is the most widely used measure for MI fidelity. The scale was originally derived from factor analysis of MI treatment sessions coded with the Motivational Interviewing Skills Code (MISC), which produced ten elements of MI practice. Correlation estimates indicated that the MITI captured 59% of the variability in the MISC. Three blind, independent coders derived reliability estimates for the MITI. Comparison of MITI scores before and after MI workshops demonstrated good sensitivity for detecting improvement in clinical practice as result of training. The MITI has since been updated (Moyers *et al*., 2010).

The MITI is aimed to be a simpler version of the MISC for use primarily in training. There are two components; global ratings of evocation, collaboration, autonomy/support, direction and empathy (measured on a 5-point Lickert scale) and behavioural counts of giving information, MI adherent utterances, MI non-adherent utterances, questions (closed and open) and reflections (simple and complex). Intra-class correlations across raters have been documented as 0.51 and 0.56 respectively, while validity has been demonstrated through a good correlation with MISC (Moyers *et al.* 2005).

5. The Motivational Interviewing Supervision and Training Scale (MISTS) (Madson *et al*., 2005)

The MISTS is a relatively new measure which correlates well with YACS scales that are relevant to MI. Again there are two components; the first comprising of behavioural counts and the second comprising a 16-item global rating of quality and effectiveness. Intra-class correlations for raters range from 0.41 to 0.81, with validity being demonstrated through its correlation with YACS (Madson *et al*., 2005).

CHALLENGES IN TESTING FIDELITY

The constraints which act to limit the likelihood of researchers engaging in the testing of fidelity of MI interventions are largely fourfold; time, money, expertise and effort. The time required to test the fidelity of a MI intervention is far greater than what would be required for, say, a blood test. Observing a video-recorded consultation and scoring it against any of the MI fidelity testing criteria will take at least as long as the treatment session, usually far longer. The Motivational Interviewing Skill Code, for example, requires the rater to make judgments on three passes through the video material. Furthermore, the rater needs to be appropriately skilled in assessing the MI intervention, usually requiring at least two-days initial and then on-going training (El-Malakh *et al*., 2012). In their manual of The Motivational Interviewing Treatment Integrity Scale, Moyers and colleagues (2010) suggest that a three-level learning programme is required for individuals who will be using the MITI to assess treatment fidelity, with intense training and assessment at each level taking on average 40 hours. That done, the rater should commit to weekly sessions of ongoing review. The use of pre-scored gold standard transcripts has been found to assist in evaluating coder competency and areas for improvement. As a result, ensuring adequate MI fidelity is likely to be costly in terms of assessment time, personnel time (requiring the time of an additional person) and recalibration where low fidelity is found. Interestingly, clinical experience (that is, being a clinician) has not predicted ease of training or eventual competence in scoring MI fidelity

(Moyers *et al*., 2010).

TECHNIQUES FOR ENHANCING FIDELITY

Techniques for enhancing testing of MI fidelity include training, sharing of experience, mentoring and involvement in established MI fidelity networks. Training an individual in MI is an obvious requirement of treatment fidelity. A pertinent question, however, is what exactly is the rater going to be trained in? It seems logical that MI fidelity ratings will be maximized by training the rater in the very behaviours listed in whatever fidelity testing instrument has been selected; training the rater to the scale. However, two problems emerge – item reliability and the validity of the items (itself an area for discussion among researchers). Most of the scale items described in this paper appear to have low inter-rater reliability, meaning that even extensive training in specific items may lead to a low fidelity score if the rater interprets the item differently from the trainer. A relevant validity question is 'if we train an individual to engage in a particular behaviour, how can we be sure we have selected the appropriate behaviour?' However, what is probably more important is the sharing of experiences through collaboration with researchers using similar MI fidelity testing techniques. This will enable individuals to seek guidance on the specific skills that comprise whichever approach is relevant to them, as well as facilitate partaking in the spirit of Motivational Interviewing through networks of individuals working in this area. The international Motivational Interviewing Network of Trainers (MINT) is one example of this (http://www.motivationalinterviewing.org/about_mint).

FIDELITY VERSUS SKILL

There are clearly issues pertaining to the relationship between MI intervention fidelity and clinician skill. Can a clinician's skill be reduced simply to the extent to which they engage in certain key behaviours? Is an effortless tally of behaviours enough or does clinical skill lie in the gaps between utterances? Is it a question of not only knowing what to say, but when not to say something? More broadly, do we know what the appropriate concepts to be measuring in MI actually are? Even Rollnick, one of the earliest proponents of MI, has suggested that the active ingredients of MI are not clear (Rollnick, 2001).

Further, is a non-adherent statement by a clinician the exact opposite of an adherent statement? Many MI fidelity assessing instruments, as demonstrated, require tallies of adherent and non-adherent behaviours. But how do we combine these? Is there an equation such that three adherent statements are cancelled out by one non-adherent statement? Or might an occasional apparently non-adherent statement sometimes be designed to create ambivalence? It seems reasonable that these and other situations could sometimes arise that challenge the ability of MI fidelity instruments to appropriately discriminate; thus affecting both reliability and validity. This is not a new question; indeed, the literature around psychological therapies has been haunted for many years by some therapists appearing to be effective regardless of what they to do (Sarris, 2011). Put another way, there appears to be a skill element conceptually separate to adherence to method that may have just as much impact on the success of a MI intervention or otherwise. While not a licence for individuals to justify whatever they do, it is a challenge for approaches strictly based on counting of specified behaviours.

Another tautology prevalent among psychological interventions concerns instances where a technique is found to be ineffective but the researchers believe this is solely because the appropriate technique was not complied with. This is another reason why it is so important to be able to demonstrate that the therapist did what they were supposed to do.

In conclusion, assessing treatment fidelity in motivational interviewing interventions is complex, challenging and vitally important.

REFERENCES

Almomani Barsky A, Coleman H. Evaluating skill acquisition in motivational interviewing: the development of an instrument to measure practice skills. *J Drug Educ* 2001; **31:** 69-82.

Carroll KM, Nich C, Sifry RL, Nuro KF, Frankforter TL, Ball SA, Fenton L, Rounsaville BJ. A general system for evaluating therapist adherence and competence in psychotherapy research in the addictions. *Drug Alcohol Depend* 2000; **57:** 225-238.

Carroll KM, Farentinos C, Ball SA, Crits-Christoph P, Libby B, Morgenstern J, Obert JL, Polcin D, Woody GE. MET meets the real world: design issues and clinical strategies in the Clinical Trials Network. *J Subst Abuse Treat* 2002; **23:** 73-80.

Corvino J, Carroll K, Nuro K, Nich C, Sifry R, Frankforter T, Ball S, Fenton L, Rounsaville B. *Yale Adherance and Competence Scale (YACS) Guidelines*. Connecticut; Yale University Psychotherapy Development Center Substance Abuse Center, 2000.

El-Mallakh P, Chlebowy DO, Wall MP, Myers JA, Cloud RN. Promoting nurse interventionist fidelity to motivational interviewing in a diabetes self-care intervention. *Res Nurs Health* 2012; **35:** 289-300.

Madson MB, Campbell TC, Barrett DE, Brondino MJ, Melchert TP. Development of the Motivational Interviewing Supervision and Training Scale. *Psychol Addict Behav* 2005; **19:** 303-310.

Miller WR, Mount KA. A small study of training in motivational interviewing. Does one workshop change clinician and patient behaviour? *Behav Cog Psychother* 2001; **29:** 457-471.

Moyers TB, Martin T, Manuel JK, Hendrickson SM, Miller WR. Assessing competence in the use of motivational interviewing. J *Subst Abuse Treat* 2005; **28:** 19-26.

Moyers TB, Martin T, Manuel JK, Miller WK, Ernst D. *Revised Global Scales: Motivational Interviewing Treatment Integrity 3.1.1*. Alberqueque, University of New Mexico Center on Alcoholism, Substance Abuse and Addictions. 2010.

Rollnick S. Comments on Dunn et al.'s "The use of brief interventions adapted from motivational interviewing across behavioral domains: a systematic review". Enthusiasm, quick fixes and premature controlled trials. *Addiction* 2001; **96:** 1769-1770.

Sarris J. Clinical depression: an evidence-based integrative complementary medicine treatment model. *Altern Ther Health Med* 2011; **17:** 26- 37.