

Report for Child Youth and Family
Ministry of Social Development



Motivational Interviewing Training Programme Evaluation

Emma Dear,

Dr. Eileen Britt

Department of Psychology

University of Canterbury

August 2014

Abstract

Motivational Interviewing (MI) is a collaborative conversation for building and strengthening motivation to change, which has been found to be effective in promoting behaviour change across a wide range of behaviours. The efficacy of MI, like other evidence-based counselling interventions, relies on the skill of the practitioner. Therefore, it is imperative that practitioners receive sufficient training in order to successfully implement the method. The present report details an evaluation of an MI training programme for staff of Child Youth and Family (CYF), Ministry of Social Development, New Zealand. Measures of MI knowledge and skill were utilised in order to evaluate the progress of the participants, and suggest that the CYF staff who attended the MI training significantly improved both their knowledge and skill in MI, although level of MI skills did not tend to be at proficiency level. Consequently further training in the form of ongoing coaching and feedback is recommended.

Ethical Approval

Ethical approval for the research was obtained from the:

- Human Ethics Committee, University of Canterbury
- Research Access Committee, Child Youth and Family , Ministry of Social Development.

Introduction

Motivational Interviewing (MI) was developed by Miller and Rollnick (1991; 2002; 2012) as a collaboration conversation for building and strengthening intrinsic motivation as a way of facilitating change. One of the beliefs of MI is that motivation to change lies within the individual and needs to be evoked in clients rather than imposed. An MI approach therefore involves the practitioner guiding the conversation to elicit intrinsic motivation from clients rather than giving advice or insight. The responsibility for change, and decision whether to change or not, rests with the client. This underlying perspective with which one practices MI has been identified as extremely important for effective practice (Miller & Rollnick, 2002). Without adopting this spirit a practitioner could succumb to what Miller and Rollnick (2002) call the “righting reflex”, or the desire to fix one’s clients.

Much of the initial work in MI was centred on substance abuse and other addictive behaviours, (Miller & Rollnick, 2002). However, MI has been widely extended to areas, including adolescents (Naar-King & Suarez, 2011), and youth offenders. For example, Stein, Monti et al. (2006) and Stein, Colby et al. (2006) investigated MI with a sample of substance abusing youth offenders and found that MI led to statistically significant reductions in: negative engagement (e.g., joking around or making fun of others) in mandated substance abuse treatment; driving under the influence of alcohol (Stein, Colby et al., 2006); and alcohol and marijuana use three months after treatment completion (Stein et al., 2011). They also reported increased positive engagement (e.g., appreciating the treatment staff and having aspirations about changing substance use or criminal behaviour), although this was not statistically significant.

There is now over 200 randomised trials of MI across different contexts and behaviours, with research generally supporting its efficacy (e.g., Lundahl, Kunz, Brownell, Tollefson & Burke, 2010). Like other evidence-based interventions, however, the efficacy of MI relies on the training of the practitioners implementing the method. Without sufficient training, effective implementation can be compromised, adversely affecting outcomes.

The present report details the evaluation of an MI training programme for Child Youth and Family (CYF) staff. Two measures were chosen to determine whether the participants improved their knowledge and skill in MI. The measures were the Motivational Interviewing Knowledge and Attitudes Test (MIKAT) which assesses MI knowledge and attitudes and the Video Assessment of Simulated Encounters–Revised (VASE-R), a test of MI skill.

Method

Procedure

Training workshops were conducted between May 2012 and July 2013. CYF staff from five regions across New Zealand (NZ) attended a 2-day MI workshop provided by Eileen Britt, a member of the Motivational Interviewing Network of Trainers (an international collective of MI trainers and researchers). Because the most recent revision of the MI spirit and structure (engaging, focussing, evoking, and planning) was published late in 2012 (Miller & Rollnick, 2012) four of the five initial workshops were provided before these changes (Miller & Rollnick, 2002). The fifth workshop was much later than the others and was carried out specifically to train newly employed staff and, whilst very similar to the original 2-day workshops, included the most recent developments in MI. Prior to the 2-day workshops, participants were invited to read two articles on MI as preparation for the workshops: Ten Things Motivational Interviewing is Not (Miler & Rollnick, 2009) and Towards a Theory of Motivational Interviewing (Miller & Rose, 2009). Participants completed the MIKAT immediately prior to and after the 2-day workshop.

The 2- day workshop consisted of video-taped demonstrations, modelling, didactic teaching along with real-play and role-playing with feedback, consistent with the recommendations for MI training (Miller & Rollnick, 2002). The content focused on the MI spirit and the skills along with emphasis on the concepts of resistance, ambivalence, and change talk. The principles of MI (express empathy, support self-efficacy, developing discrepancy, and roll with resistance) were also included,

along with research presented to demonstrate the effectiveness of MI and how it works. There was also an introduction to MI strategies such as the importance of confidence rulers.

More senior CYF staff participated in the 1-day advanced training provided by Eileen Britt. The amount of time between initial and advanced trainings varied for each location. Participants were updated on the revised spirit and structure of MI (Miller & Rollnick, 2012) during this workshop and completed the VASE-R both before and after this workshop.

The 1-day advanced workshop consisted of video-taped demonstrations, modelling, didactic teaching, and role-playing with feedback, consistent with Miller and Rollnick's (2012) recommendations for MI training. This workshop included a re-cap on the concepts of ambivalence and change talk and the importance of the role of these concepts in MI. There was also a reminder of the spirit and its importance. This training focused on the practice of MI skills, in particular reflective listening as a means of generating, evoking, and responding to change talk.

Measures

MIKAT. Developed by Leffingwell (2006) a test of knowledge and attitudes consistent with MI and its spirit, and consists of two parts:

1. A quiz comprised of true-false questions about 10 addiction myths and four MI-consistent attitudes and assumptions.
2. A counselling behaviours checklist including five behaviours prescribed for an MI approach.

A summary score can be generated by calculating the number of correct items. Additionally, separate scores for both attitudes and knowledge can be calculated. Furthermore these sub-scales themselves can be divided into subcomponents providing information on the amount of MI-

consistent attitudes versus addition myths answered correctly in the true–false component, along with the amount of correct versus incorrect behaviours selected in the checklist.

For the purposes of the present training, modifications were made to the true–false component of the MIKAT. The original true–false component is comprised of statements identifying clients as substance abusers. Because the clients of the CYF staff were youth who may or may not have had a problem with substance use, subtle changes were made to the vocabulary (but not the meaning) of some items to better portray clients in general.

VASE-R. Provides a standardised MI challenge and assesses a variety of MI skills (Rosengren, Hartzler, Baer, Wells & Dunn, 2008). This measure consists of three video-based scenarios with different substance abusing clients played by actors. After each client there are questions that require written responses which are scored. There are 18 questions that cover skills within five subscales: reflective listening (RL), responding to resistance (RR), summarising (S), eliciting change talk (ECT), and developing discrepancy (DD).

Of the 18 items, 15 are of a free-response format with the remaining three being multiple choice (Rosengren et al., 2008). Items are scored using a 3-point system. For RL, RR, and ECT, a 0 is given if the response is likely to evoke resistance or is confrontational; if the response is an inaccurate representation of the client’s speech or neutral, a 1 is given; if there is an accurate representation of the intended MI skill and reflection of the client’s speech, a 2 is given. For S, a 0 is given if the response is again confrontational or likely to evoke resistance, but also if it does not include multiple ideas. A 1 is given if the response has multiple ideas, however only includes either ambivalence or client change talk. It is only when the response includes multiple ideas along with both ambivalence and client change talk that a 2 is given. Finally for the DD subscale, a 0 is given if an incorrect option is chosen and paired with a rationale not consistent with MI; a 1 given if again an incorrect option is chosen, however the rationale is MI consistent; for 2 points the correct option

must be chosen. The collated points yield a full-scale score that can range from 0–36. Rosengren et al. (2005) suggest that an overall score of 27 on the VASE-R indicates general MI proficiency.

A NZ edition of the VASE-R has been developed using actors with NZ accents and vocabulary to portray clients (Hall, McMaster & Associates Limited, 2012). The content of the DVD is analogous to that of the original measure and the original scoring form, manual and answer sheet are used (Rosengren et al., 2005; 2009).

The VASE-R Administration and Scoring Manual (Rosengren et al., 2005) states that 90% inter-rater agreement is recommended. The rater had no experience with the VASE-R prior to the present research, therefore training was provided by Eileen Britt, which comprised reviewing the general concepts of MI along with the various skills required to effectively perform the method. The VASE-R was then completed by the rater and the responses were discussed whilst reviewing the VASE-R manual. The VASE-R was also completed by a third party and independently scored to assess the level of agreement. Once the desired level of agreement was achieved, the rater was permitted to begin rating the participants' VASE-Rs. Additionally, a subset of 19 (20%) of the VASE-Rs were also scored by the rater and Eileen Britt to ensure an adequate level of inter-rater reliability was maintained.

Data Analysis

The primary analyses focussed on the effects the training had on the MI knowledge and skill attainment of the participants. Additional analyses investigated the pattern of these training effects across the different workshop locations.

MIKAT. Because participants names were not on the answer sheets, it was not possible to pair scores or determine the ownership of missing tests, therefore a nonparametric test was deemed appropriate. Mann-Whitney U tests were performed to determine whether there was a significant difference between pre- and post-training scores from the 2-day workshop. A power

analysis determined that for 95% power, a sample size of $n = 110$ pre- and post-training was required assuming a medium effect size ($d = .50$).

A Pearson correlation was performed using the total MIKAT scores to determine the relationship between the pre- and post-training results. Agreement between the pre- and post-total scores was then measured using the Gower coefficient of agreement (Gower, 1971). As the workshops were implemented across several locations over an extended period of time (approximately 12 months), high agreement on the Gower coefficient would indicate that all groups improved similarly on the MIKAT regardless of location or time of workshop.

The Gower index was computed using the “Gower” computer software version 1.1 (www.pbarrett.net/software.html). Because there is no obvious hypothetical sampling distribution for the Gower coefficient, a bootstrap procedure (www.pbarrett.net/Bootstrap/Bootstrap.html) was employed to compute credibility intervals (the interval where it is expected to observe 95% of all coefficients computed using the same sample size, number-type and same minimum and maximum possible data range as that for the observed coefficient).

VASE-R. Prior to the analyses of training effects, inter-rater reliability was assessed. Two-way mixed, absolute agreement, average measures intra-class correlations (ICCs) were used to assess the degree to which the raters demonstrated agreement in their overall and subscale ratings of the VASE-R.

To determine whether the participants improved on MI skill as a result of training, a paired samples t-test was performed to identify whether there was a significant difference in scores between pre- and post-training, and also whether the participants had reached proficiency in MI. A power analysis determined that for 95% power, a sample size of $n = 54$ was required assuming a medium effect size ($d = .50$).

Pearson correlation was performed, as with the MIKAT to determine the relationship between the pre- and post-training results. The correlation was then disattenuated to account for measurement error and agreement between the pre- and post-total scores was measured using the Gower coefficient of agreement (Gower, 1971) and the bootstrapping procedure to obtain significance. Again, a high Gower coefficient would indicate similar improvement across all groups regardless of location of the workshops.

Results

MIKAT

Additional prior analyses that were performed to evaluate the MIKAT itself (Dear, 2014) suggested problems with the validity of the MI-consistent attitudes subscale. This subscale was therefore excluded from further analyses. Thus that MIKAT no longer measured MI attitudes and instead could be considered a measure of MI knowledge only. Additionally, the earlier analysis also identified that the MIKAT was measuring different aspects of MI knowledge. It was therefore decided to analyse the measure as having two separate components: the myth scores and correct checklist choice scores; and the incorrect checklist choice scores.

With the exclusion of the MI-consistent attitudes subscale, the total MIKAT score was out of 15. The means and standard deviations for this total score increased from pre-training (8.73 ± 4.01) to post-training (11.56 ± 2.71). Additionally, the incorrect choice scores decreased from pre- to post-training ($2.12 \pm 2.38 - 1.58 \pm 2.45$).

A Mann-Whitney U test was run to determine whether there was a significant difference in the total and incorrect checklist choice score of the MIKAT from pre- to post-training. Eighteen participants had incomplete scores on the MIKAT, eleven did not complete a pre-test and seven did not complete a post-test. According to Little's MCAR test (Little, 1988) the missing data were not missing completely at random.

With a sample size of $n = 74$ for the pre-training and $n = 78$ for post-training the analyses had 85% power assuming a medium effect size ($d = .50$). The total MIKAT mean ranks significantly increased from pre- (56.66) to post-training (93.43), $U = 1,565.50$, $z = -2.30$, $p < .001$, $d = .83$. With regards to the incorrect checklist choice scores, a significant decrease was observed in the mean ranks from pre- (84.60) to post-training (68.81), $U = 2,286.5$, $z = -2.30$, $p = .02$, $d = .22$. Collectively, these results indicate that from pre- to post- the 2-day workshop training, the participants' MI knowledge significantly increased.

Pearson correlation revealed that there was no relationship between the total MIKAT pre- and post-training scores ($r = .02$, $p = .867$). Additionally, there was a statistically significant high coefficient of agreement ($.76$, $p < .001$) indicating that, irrespective of location, the changes in scores from pre- to post-training were 76% similar. The prior Mann-Whitney U test determined that this was a positive change, i.e., scores increased from pre- to post-training.

VASE-R

Inter-rater reliability. The ICCs for the total VASE-R scores was in the excellent range ($ICC = .935$), indicating a high level of agreement. There was some variation amongst the subscales, with the strongest reliability being for the DD scale ($ICC = 1.000$). This was closely followed by RL ($ICC = .952$), RR ($ICC = .868$), ECT ($ICC = .820$), and S ($ICC = .762$). Despite the variation, all remained in the excellent range ($.75 - 1.00$, Cicchetti, 1994).

Sensitivity to training effects. Again, additional analyses were performed to investigate the VASE-R (Dear, 2014). During these analyses it was identified that there was much overlap in what was being measured by the five subscales comprising the VASE-R. This indicated that analysis of the VASE-R as a single measure was most appropriate and therefore the measure was analysed using only the total score.

There was a significant amount of missing data among the sample: eight participants had incomplete scores; seven did not complete a post-test and one did not complete a pre-test. The data were not missing at random according to Little's MCAR test (Little, 1988). With a sample size of $n = 38$ the analyses had 85% power to detect a medium effect size ($d = .50$). A preliminary assumption check indicated that data were normally distributed as assessed by a Shapiro Wilk test ($p = .46$) and that no outliers were detected.

Participants improved their scores from pre- to post the 1-day workshop training (14.93 ± 6.82 – 19.05 ± 7.19), and this increase was statistically significant ($p < .001$, $d = 0.59$). With the mean post-training score of 19.05, however, it appears that, as a whole participants, did not reach general MI proficiency on the measure despite the significant increase in VASE-R scores. Five of the 46 participants attained a score at or above the MI proficiency score of 27 suggested by Rosengren et al. (2005).

There was a high correlation between the total VASE-R pre- and post-training scores ($r = .83$). The correlation was then disattenuated for measurement error, increasing the value to demonstrate perfect linearity ($r_c = 1.009$). There was a significantly high Gower coefficient of agreement ($.82$, $p < .001$) indicating that, irrespective of location, the changes in scores from pre- to post-training were 82% similar. The prior t-tests indicated that these changes were in a positive direction (i.e., scores increased from pre- to post-training).

Discussion

Because MI has been shown to be effective in promoting behaviour change, the importance of effective training is imperative to ensure it is implemented in a manner likely to produce positive outcomes. Two MI measures, the MIKAT and the VASE-R, were used to evaluate the CYF MI training programme. Participants' scores on both measures statistically significantly increased from pre- to post-training which indicated that the training programme significantly improved participant MI

knowledge and skill. Additionally, score increases across the five training groups (different locations) were relatively consistent for both measures: 76% of participants who completed the MIKAT had similar increases in their scores and 82% of VASE-R participants' scores increased by 10 points or less. Collectively these scores indicate that not only did the participants statistically significantly improve on both MI knowledge and skill as a result of the training, but a significant percentage of participants' on both measures increased similarly across all the training groups. Whilst the study did not have a control group, the fact that similar results were obtained with each repeated workshop, at different locations, and over time, suggests that the improvements in MI knowledge and skill demonstrated post workshop training are likely to have been the result of the workshops rather than any other extraneous factor.

However, the VASE-R scores indicate that the majority of participants did not reach MI proficiency. This suggests that more training is needed to assist with further MI skill development. This finding is consistent with research on MI training (Miller and Rollnick , 2012; Miller, Yahne, Moyers, Martineez & Pirritano, 2004) which suggests that ongoing feedback and coaching beyond workshop based training is needed to develop proficiency in MI and to facilitate the transfer of MI into the work place environment.

References

- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*(4), 284–290 .
- Gower, J. C. (1971). A general coefficient of similarity and some of its properties. *Biometrics, 8*57–871.
- Hall, McMaster & Associates Limited (Author) (2012) *Motivational Interviewing DVD* [learning resource], Auckland: Sticky Fingers.
- Leffingwell, T. R. (2006). Motivational Interviewing Knowledge and Attitudes Test (MIKAT) for evaluation of training outcomes. *MINUET, 13*, 10–11.
- Little, R. J. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association, 83*(404), 1198–1202.
- 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.
- Miller, W., & Rollnick, S. (1991). *Motivational Interviewing: preparing people to change addictive behaviour*. New York: Guilford Press.
- Miller, W. R., & Rollnick, S (2002). *Motivational Interviewing: Preparing People for Change*. New York; London : Guilford Press.
- Miller, W. R., & Rollnick, S. (2009). Ten things that motivational interviewing is not. *Behavioural and Cognitive Psychotherapy, 37*(2), 129–140.
- Miller, W. R., & Rollnick, S. (2012). *Motivational Interviewing: Helping people change*. New York; London: Guilford Press.
- Miller, W. R., & Rose, G. S. (2009). Toward a theory of motivational interviewing. *The American Psychologist, 64*(6), 527–537.

- Miller, W.R., Yahne, C.E., Moyers, T.B., Martinez, J. & Pirritano, M. (2004). A randomized trial of methods to help clinicians learn motivational interviewing. *Journal of Consulting and Clinical Psychology, 72*, 1050-1062.
- Naar-King, S. & Suarez, M. (2011). *Motivational Interviewing with adolescents and young adults*. New York: Guilford Press.
- Rosengren, D. B., Baer, J. S., Hartzler, B., Dunn, C. W., & Wells, E. A., Ogle, R. (2005, revised 2009). Video assessment of simulated encounters (VASE-R): Administration and scoring manual. *University of Washington, Seattle, WA*.
- Rosengren, D. B., Hartzler, B., Baer, J. S., Wells, E. A., & Dunn, C. W. (2008). The video assessment of simulated encounters-revised (VASE-R): Reliability and validity of a revised measure of motivational interviewing skills. *Drug and Alcohol Dependence, 97*(1), 130–138.
- Stein, L. A., Colby, S. M., Barnett, N. P., Monti, P. M., Golembeske, C., & Lebeau-Craven, R. (2006). Effects of motivational interviewing for incarcerated adolescents on driving under the influence after release. *The American Journal on Addictions, 15*(1), 50–57.
- Stein, L., Lebeau, R., Colby, S. M., Barnett, N. P., Golembeske, C., & Monti, P. M. (2011). Motivational interviewing for incarcerated adolescents: effects of depressive symptoms on reducing alcohol and marijuana use after release. *Journal of Studies on Alcohol and Drugs, 72*(3), 497–506.
- Stein, L., Monti, P. M., Colby, S. M., Barnett, N. P., Golembeske, C., Lebeau-Craven, R., & Miranda, R. (2006). Enhancing substance abuse treatment engagement in incarcerated adolescents. *Psychological Services, 3*(1), 25–34.