MOTIVATIONAL INTERVIEWING: A REVIEW OF CODING SYSTEMS

Jennifer K. Manuel, PhD[1], Theresa B. Moyers, PhD[2], & Jon M. Houck, MS, MA[2]
[1]Department of Psychiatry, University of California, San Francisco;
[2]Department of Psychology, University of New Mexico

Introduction

As the field of substance use continues to focus on the dissemination and implementation of evidence-based practice, questions arise on the most efficient and efficacious ways of training providers in evidence-based treatments. Training in motivational interviewing (MI) has been a particular focus among researchers, presumably because of its evidence-base among a variety of health behaviors. The evaluation of MI sessions using behavioral coding systems has increased substantially in recent years. Coding systems are used for a variety of purposes: to measure treatment fidelity to MI principles, to examine the mechanisms of change in treatment sessions, and to better understand the unique contributions of therapist and client language in session. As such, there are a number of coding systems available to researchers and clinicians.

Methods

Behavioral coding systems and systematic measures of MI treatment adherence and fidelity were identified from a recent review article and via a literature search using Medline and PsycINFO databases. A total of 11 published behavioral coding systems for motivational interviewing skills were included in the current review.

Table 1. Review of Coding Systems

Coding System	Overview	Psychometrics
Behavior Change Counseling Index (BECCI; Lane et al., 2005)	The BECCI is an 11-item questionnaire that measures health behavior change and MI skills in physician-patient interactions. Items are rated on a 0-4 scale. Sample items include, "Practitioner encourages patient to talk about change" and "uses empathic listening statements when patient talks abut the topic."	This measure has good construct validity (cronbach alpha = .71 and .63) as assessed in two samples. Inter-rater reliability was assessed using correlation coefficients and the BECCI demonstrated good reliability (r=.079 and r = .93) and good intra-rater reliability (r=.60 – r= .90). The measure also shows sensitivity to training.
Client Evaluation of Motivational Interview (CEMI; Madson et al., 2009)	35-item questionnaire completed by clients. Assesses counselors MI spirit and MI skills. Items are rated on 1-4 scale. Example questions, "Helps you feel confident in your ability to change your behavior." Takes 10 minutes for client to complete.	Good internal consistency in a study (N=60) of homeless, mostly African American men in α=.83.
Independent Tape Rating Scale (ITRS: Martino et al., 2008)	39-item behavioral coding system measuring 1) Fundamental MI skills (open questions, affirmations), 2) Advanced MI skills (develop discrepancy), 3) MI Inconsistent behaviors (giving advice or direction), 4) General substance abuse counseling.	28 of 30 items showed good-excellent ICCs with fair reliability for the other two items.
MI Process Code (Barsky & Coleman, 2001)	Videotaped sessions are rated on a 1-5 scale for 1) Functional Skills (n=13) and 2) Dysfunctional Skills (n=12). Sample items from the Functional Skills scale include, "Expresses empathy" and "Helps client identify barriers to change." Dysfunctional Skills scale items include, "Argues or debates with client" and "Label client."	In the Barksy & Coleman, (2001) study, eight sessions with simulated patients were coded. Inter-observer reliability was calculated by the following equation: Number of agreements/(number of agreements + disagreements). Coders achieved 51% for Functional skills items and 75% for Dysfunctional skills.
MI Skills Code (MISC; Moyers, Martin, Catley, Harris, & Ahluwalia, 2003)	The MISC is coded in three passes, during which the session is listened to in its entirety. In pass 1, the global scores are measured for the therapist (Acceptance, Egalitarianism, Empathy, Genuineness, Warmth and Spirit), client (Affect, Cooperation, Disclosure, Engagement) and therapist –client interactions (Collaboration, Benefit). In Pass 2, therapist behaviors (e.g., Direct, Rephrase, Summary) and client behaviors (e.g., Ask, Change talk, Resisting Change) are tallied. In the third pass, the percentage of therapist and client talk time during the session is calculated.	Global reliability ICC ratings ranged from .25 - .79. Client cooperation, client affect and therapist acceptance fell into the "poor" range while therapist empathy, spirit, and client engagement were in the "excellent" range. Behavior count reliability ratings were also varied with ICCs ranging from .00 – 1.00.
MI Supervision and Training Scale (Madson et al., 2005).	The MISTS is comprised of 1) therapist behavioral counts and 2) 16-item global ratings. Examples of behavior counts include open/closed questions, simple/complex reflections, affirmations, summaries, and failure to elicit or reinforce client change talk. These behaviors are tallied throughout the session. Global ratings were rated on a 1-4 scale and measured therapist listening skills, MI spirit skills, and general therapist ratings.	Madson et al. (2005) found that ICCs for individual items ranged from .41 - .81. The MISTS total scale was significantly correlated with the Assess and Support measures from the YACS. The MISTS total scale was uncorrelated with the Clinical Management and Twelve-Step facilitation scales and was correlated with the Cognitive-Behavioral Scale from the YACS.
Motivational Interviewing Treatment Integrity (MITI; Moyers et al., 2005)	The MITI is derived from the MISC using an exploratory factor analysis. The original MITI (Moyers et al., 2005) was comprised of 2 global measures and 7 therapist behavior counts. Global measures (Empathy, Spirit) are rated on a 1-7 Likert scale. Behavior counts (open questions, closed questions, simple and complex reflections, general information, MI-Adherent and MI-Nonadherent) are tallies of each instance of a therapist behavior in the session. The MITI has been revised and now all global ratings are measured on a 1-5 scale. The Spirit rating is now parsed into three ratings: Autonomy/Support, Collaboration, and Evocation. A Direction global rating has been added.	Moyers et al. (2005) reported ICCs of .52 for Empathy and .58 for Spirit. Behavior count ICCs ranged from .58 - .97. MITI scores from pre and post MI training workshops significantly differed, indicating that the MITI is sensitive enough to detect differences in MI skills.
Sequential Code for Process Exchanges (SCOPE; Moyers et al., 2007)	The SCOPE was developed from the MISC and is an in-depth therapist and client coding system. Thirty therapist and 16 client codes are tallied. Therapist behaviors are categorized into: MI Consistent, MI Inconsistent or Other. Client language is categorized into Change Talk or Counter Change Talk. This coding system allows for sequential examination of therapist and client behaviors.	Reliability estimates for the SCOPE indicate that the categories were reliably differentiated. Cohen's k was .66 for MI Consistent and .68 for MI consistent. Client change talk was .70 and counter change talk was .67.
Stimulated Client Interview Rating Scale (SCIRS; Arthur 1999)	This 39-items scale measures nursing students' "basic humanistic communication skills" and basic MI skills with simulated patients. There are two sections (Communication skills and Menu of Strategies). All items are rated on a 1-3 scale with 1 = not done, 2 = done, and 3 = done well. Sample Communication skill items include: "Sits squarely, demonstrates reflective listening". Sample Menu of Strategies items include: "Chooses right moment to ask permission, uses simple questions to ask about a typical day, describes a standard drink."	In the Arthur (1999) study, video sessions were rated by the researcher, simulated client and student nurse. Total scores were compared among the three raters. There was a significant positive correlation (r=.297, p < .001) between student nurse and researcher scores. Simulated clients and researcher scores were weakly correlated (r=.108, p < .01) and student and client scores were uncorrelated (r=.108, p > .01). Cronbach alpha was greater than .90 for each rating (client, student and researcher).
Video Assessment of Simulated Encounters-revised (VASE-R; Rosengren et al., 2008)	The VASE-R is an evaluation system that can be used with groups or individuals. Respondents view videotaped simulated client vignettes and are then asked to write brief answers. Questions are either open-ended or multiple choice that ask respondents to provide a rationale for their choice. Answers are measured from 0-3 with higher scores reflecting MI skills. The VASE-R has five subscales: 1) Reflective Listening; 2) Responding to Resistance; 3) Summaries of client ambivalence and change talk; 4) Eliciting change talk; 5) Develop discrepancy.	ICCs ranged from .41 - .92 and .50 - .96 in two studies of the VASE-R (Rosengren et al., 2008). The measure showed good internal consistency (cronbach alpha = .85). The VASE-R demonstrated sensitivity to MI training effects and displayed good concurrent validity with the MITI.
Yale Adherence and Competence Scale (YACS; Carroll et al., 2000)	A 55-item behavioral rating scale that can be used to assess general therapist behaviors and specific therapist behaviors (for Clinical Management, Twelve-Step Facilitation, and Cognitive-Behavioral Therapy). Each item is rated on a 1-5 Likert scale for adherence and competence, allowing raters to differentiate therapists who may deliver an intervention in a less skillful manner.	All scales were reliable among coders in their study with ICCs ranging from .80-.95 for adherence ratings and .71-.97 for competence ratings. Among individual items reliability was varied, with ICCs ranging from .28 - .84 among adherence items and .06 - .81 among competence items. A confirmatory factor analysis indicated that the six scales met goodness-of-fit criteria. The three treatment scales were negatively correlated.

Supported by NIDA T32 DA007250, R21DA025135, R01DA021227

Results

Table 1 describes the 11 coding systems reviewed. The coding systems vary greatly in the level of detail provided. Most coding systems (N=9) focus on the therapist or practitioner. Only the MISC and the SCOPE specifically categorize client language. This additional client data is not without a cost, as both the MISC and SCOPE are time-consuming coding systems.

Information Provided

Some coding systems are in a checklist format (e.g., BECCI), with coders completing Likert scale items for the various domains of MI skills. Other coding systems use behavior count tallies with coders counting each instance of a behavior. Only the SCOPE provides sequential data between therapist and client language. This is the only coding system that can indicate the precise relationships between therapist and client behaviors. The MISC also categorizes client language but allows only for overall correlations between therapist and client within-session behaviors. The YACS is the only coding system that assesses specific skills related to MI, Clinical Management, and CBT.

Coder Training

In many of the studies, coders were highly educated and experienced in substance abuse counseling. Many research studies provided coders with in-depth training in the coding systems. Despite this extensive training and experience, many coding systems displayed poor inter-rater reliability for a number of items.

Conclusions

As MI continues to gain popularity in research and clinical settings, coding systems are becoming more widespread to measure practitioner treatment fidelity to MI as well as to understand the underlying mechanisms in MI treatment sessions. The selection of an appropriate coding system is an important consideration and should be based upon what information is needed from the coding (e.g., treatment fidelity, measure of client language) and the time and financial resources available.

References

Arthur, D. (2009). Assessing nursing students' basic communication and interviewing skills: the development and testing of a rating scale. *Journal of Advanced Nursing 29*(3), 658-665.

Barsky, A. & Coleman, H. (2001). Evaluating skill acquisition in motivational interviewing: The development of an instrument to measure practice skills. *Journal of Drug Education, 31*(1), 69-82.

Carroll, K.M., Nich, C., Sifry, R.L., Nuro, K.F., Frankforter, T.L., Ball, S.A., Fenton, L., & Rounsaville, B.J. (2000). A general system for evaluating therapist adherence and competence in psychotherapy research in the addictions. *Drug and Alcohol Dependence, 57*, 225-238.

Lane, C., Huws-Thomas, M., Hood, K., Rollnick, S., Edwards, K. & Robling, M. (2005). Measuring adaptations of motivational interviewing: the development and validation of the behavior change counseling index (BECCI). *Patient Education and Counseling, 56*(2), 166-1873.

Madson, M.B., Bullock, E.E., Speed, A.C., & Hodges, S.A. (2009). Development of the client evaluation of motivational interviewing. *MINT Bulletin, 15*(1), 6-8.

Madson, M.B., Campbell, T.C., Barrett, D.E., Brondino, M.J., & Melchert, T.P. (2005). Development of the motivational interviewing supervision and training scale. *Psychology of Addictive Behaviors, 19*(3), 303-310.

Martino, S., Ball, S.A., Nich, C., Frankforter, B.A. & Carroll, K.M. (2009). Informal discussions in substance abuse treatment sessions. *Journal of Substance Abuse Treatment, 36*(4), 366-375.

Moyers, T.B., Martin, T., Catley, D., Harris, K.J., & Ahluwalia, J.S. (2003). Assessing the integrity of motivational interviewing interventions: reliability of the motivational interviewing skills code. *Behavioural and Cognitive Psychotherapy, 31*, 177-184.

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

Rosengren, D.B., Hartzler, B., Baer, J.S., Wells, E.A., & Dunn, C. (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*(1-2), 130-138.