Journal of Psychosomatic Research xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Journal of Psychosomatic Research

Selecting an optimal instrument to identify active ingredients of the motivational interviewing-process

Jos Dobber [a,*], Berno van Meijel [b,c,e], Emile Barkhof [d], Wilma Scholte op Reimer [a], Corine Latour [a], Ron Peters [f], Don Linszen [d]

[a] Amsterdam School of Health Professions, Amsterdam University of Applied Sciences, Amsterdam, the Netherlands
[b] Inholland University of Applied Sciences, Department of Health, Sports & Welfare, Cluster Nursing, Amsterdam, the Netherlands
[c] Parnassia Psychiatric Institute, Parnassia Academy, The Hague, the Netherlands
[d] Department of Psychiatry, Academic Medical Center, Amsterdam, the Netherlands
[e] VU University Medical Center, Department of Psychiatry, EGO Institute for Health and Care Research, Amsterdam, the Netherlands
[f] Department of Cardiology, Academic Medical Center, Amsterdam, the Netherlands

ARTICLE INFO

Article history:
Received 24 May 2014
Received in revised form 8 October 2014
Accepted 20 October 2014
Available online xxxx

Keywords:
Motivational interviewing
Therapy process
Sequential coding

ABSTRACT

Objective: Motivational Interviewing (MI) can effectively stimulate motivation for health behavior change, but the active ingredients of MI are not well known. To help clinicians further stimulate motivation, they need to know the active ingredients of MI. A psychometrically sound instrument is required to identify those ingredients. The purpose of this study is to describe and evaluate the capability of existing instruments to reliably measure one or more potential active ingredients in the MI process between clients and MI-therapists.
Methods: We systematically searched MedLine, Embase, Cinahl, PsycInfo, Cochrane Central, specialised websites and reference lists of selected articles.
Results: We found 406 papers, 60 papers were retrieved for further evaluation, based on prespecified criteria. Seventeen instruments that were specifically designed to measure MI or aspects of MI were identified. Fifteen papers met all inclusion criteria, and reported on seven instruments that assess potential active ingredients of the interactive MI process. The capability of these instruments to measure potential active ingredients in detail and as a part of the interactive MI process varies considerably. Three of these instruments measure one or more potential active ingredients in a reliable and valid way.
Conclusion: To identify the potential active ingredients in the interactive MI process, a combination of the SCOPE (which measures potential technical active ingredients) and the GROMIT or the global ratings of the MISC2 (to measure potential relational ingredients) seems favourable.

Introduction

Currently Motivational Interviewing (MI) is applied in a number of target populations and problem areas and benefits from an increasing popularity. It addresses a range of behaviors, such as reducing substance abuse, diet and exercise, and other lifestyle outcomes [1–6]. Evidence suggests that MI is effective, especially in substance use disorders [1–5]. However, questions such as "How does MI work?" and "What are the active ingredients of MI?" remain unanswered [1,3,5,7].

MI is "a collaborative counseling style for strengthening a person's own motivation and commitment to change" ([8], p234). It pays particular attention to the language of change, also called "change talk" (favoring change: e.g. "I probably should quit smoking") and "sustain talk" (favoring not changing: e.g. "I don't think I can quit"), which refers to statements in which the client expresses some kind of motivation for change. MI is a complex behavioral intervention [7,8], and MI sessions are complex processes of therapist utterances influencing client utterances and vice versa, in which the therapist continually makes choices in MI techniques and strategies. Through these techniques and strategies, the therapist elicits the client's own good reasons for change, discussed within a good client-therapist relationship, and by this the active ingredients of MI are applied in the therapeutic process. These active ingredients are "the key therapist strategies that facilitate positive change" ([9], p860). For MI, however, the active ingredients are not well known, although there are some indications for potential active ingredients from research (e.g. [10]). Currently, the ingredients of the MI-process are derived from MI-theory [7,11]. If we can measure the MI-process with a focus on these (potential) active ingredients, we may obtain a better insight in the actual active ingredients within the MI-process and how they influence the patient's behavior. For this we need an instrument that measures the MI-process in a valid and reliable

* Corresponding author at: Amsterdam School of Health Professions, Amsterdam University of Applied Sciences, Tafelbergweg 51, 1105 BD Amsterdam, the Netherlands.
E-mail address: j.t.p.dobber@hva.nl (J. Dobber).

http://dx.doi.org/10.1016/j.jpsychores.2014.10.010
0022-3999/

way, meaning that the instrument should represent these potential active ingredients.

The aim of the current literature review is to describe and evaluate the capability of existing instruments to measure one or more potential active ingredients of the interactive MI-process between clients and MI therapists in a valid and reliable way. Such an instrument should measure at least a part of the potential active ingredients of the MI-process.

Conform Miller & Rose [11], we distinguish a relational and a technical category of potential active ingredients. In the "relational category" appreciation of the client-therapist-relationship, well-timed and skillfully performed empathic understanding and MI-Spirit (a composition of partnership, acceptance, and compassion), are associated with better outcomes [10–12]. The "technical category" comprises the use of MI-techniques and strategies to evoke client change talk. MI-consistent behavior is associated with more client change talk, while MI-inconsistent behavior is associated with more client sustain talk. And these client expressions are associated with treatment outcomes [8,10]. Amrhein et al. [13] found that client commitment statements predicted the effect of MI on drug use outcomes. Other client expressions, such as statements about reasons for change, and average strength of ability statements, may also be associated with improved outcomes [10,14–17]. These studies indicate that in MI client change talk may be related to "processes occurring within the client, the mechanism of change" ([9], p. 860). So, to promote change, the therapist employs the active ingredients to stimulate the mechanisms of change.

Although the processes underlying MI and its active ingredients remain unclear, some of the potential active ingredients relate to therapist behavior or to client-change talk.

Consequently, for process research, suitable instruments show (a) which relational and/or technical ingredients the therapist employs, and preferably also when and how the therapist uses these ingredients, *and/or* (b) the client motivational process, made visible through the client change talk or sustain talk. The instrument or the combination of instruments should enable the study of the effects of the therapist behavior in detail, by evaluating its immediate effect on the client behavior. To study this interactive MI-process, the order of the therapist-client interaction must be maintained as much as possible to bring into focus the interactive process. If we can identify the active ingredients of MI, clinicians will be able to purposefully apply these active ingredients, which will enhance the effectiveness of MI.

In the current literature review we will discuss the potential of the available instruments to measure (a part of) one or more potential active ingredients of the interactive MI-process. We will also evaluate the psychometric properties of these instruments.

Methods

Literature search

We searched computerized databases (MedLine, Embase, Cinahl, PsycInfo, Cochrane Central), with the following search string, using free text search terms: ((motivation OR motivational) AND (interview OR interviewing) OR (motivational interviewing)) AND (intervention fidelity OR skill OR evaluation) AND (validity OR reliability). The searches covered the period from 1990 to December 2013. No additional limits were used. This search included the bibliographies on www.motivationalinterview.org. We also searched for relevant cross-references in the reference lists of the selected articles.

Selection and quality assessment

It may be possible that an instrument or a combination of instruments jointly disclose the interactive process. The instrument must provide sufficient information to allow inferences on therapist behavior and strategies, and their effect on the client. All kinds of existent MI-instruments (e.g. training tools, research tools, proficiency measurement tools) may be suitable to contribute to this, on the condition that the instrument measures a potential active ingredient and/or its effect on the client in a valid and reliable (preferably expressed in Intraclass Correlation Coefficient/ICC or Kappa) way. Also, the measurement should be detailed enough to gain insight in the interactive process. Coding systems that divide therapist and/or client behavior in only two categories each (e.g. elicited change talk: yes/no), are considered to offer not enough information for this purpose. So we used the following inclusion criteria to select instruments for this review. (1) The instrument specifically addresses measuring the execution of MI, (2) the instrument brings into focus one or more potential active ingredients in the MI-process and/or their effect on client behavior, (3) the measurements are based on observations, and (4) the instrument collects detailed information.

Two researchers independently selected the articles based on prespecified criteria (first selection on title and abstract, second selection on full text) and each read the full text of the selected articles to perform the quality assessment. In case of disagreement on the data, the text of the original paper was checked. The quality assessment focused on the procedures, as described in the articles that reported on the studies, to assess the risk of bias (RoB) that may have occurred in the process of reliability-testing of the instrument. Since we did not find a suitable checklist to assess the RoB in the development of instruments for complex behavioral interventions, we developed a structured assessment form that all researchers used for the quality assessment. This form is based on the assumptions that (1) the reliability sample should be randomly chosen, and (2) big enough to avoid selection bias. Also, to avoid information bias, (3) the coders should be trained well enough to be able to code this complex behavioral intervention in a reliable way. Finally, also to avoid information bias, (4) to maintain the acquired coding skills, and to keep coding reliably, supervision or regular coding meetings are necessary. Hence, our structured RoB-form assessed (a) the sampling method (random or nonrandom), (b) the size of the reliability sample (the proportion of the sessions that was used to measure the inter-rater agreement), (c) the duration of the coder training (number of hours), and (d) the existence of ongoing supervision/coder meetings during the coding period.

Data extraction

The same two researchers each independently extracted the data from the selected studies and from the instruments, via a structured data extraction form. The collected data of interest included the goal of the instrument, the ingredients that are measured, the method(s) of measuring (e.g. Likert scales, behavior counts), and all information on reliability and validity measures.

Level of detail

We categorized the instruments in two categories to differentiate in level of detail: 1. instruments collecting information with a low level of detail (dividing client and/or therapist behavior in two broad categories each), 2. instruments collecting detailed information (dividing client and/or therapist behavior in three or more categories each).

Results

The systematic literature search identified 406 potentially relevant papers. Many of these papers were not on MI, or were RCTs in which the process of MI had not been measured. Sixty papers were retrieved in full text for further evaluation, revealing seventeen instruments that were specifically designed to measure MI or aspects of MI (Table 1). One instrument didn't meet the first inclusion criterion because it measures behavior change counseling (which does not strategically elicit change talk and develop discrepancy) instead of MI [19]. Three instruments are coding responses to scenarios, vignettes or simulated patients [20,32,39,40], and did not meet criterion 2. Two instruments measure through client opinion [22,37], and not through observations (criterion 3), and three instruments collect information with a low level of detail [21,23,36] (criterion 4). Finally, one instrument incorporates an other instrument (the MITI) to measure the MI-elements [35]. The eleven papers reporting on these ten instruments were excluded.

Please cite this article as: Dobber J, et al, Selecting an optimal instrument to identify active ingredients of the motivational interviewing-process, J Psychosom Res (2014), http://dx.doi.org/10.1016/j.jpsychores.2014.10.010

Fifteen papers reporting on the seven remaining instruments, met our inclusion criteria, and were included in this study (Fig. 1). Each of these instruments may measure at least one potential active ingredient, and may contribute to measure this ingredient in the interactive MI-process.

Quality/risk of bias

Table 2 presents the RoB of the included studies. In general there was a low RoB for sampling method, duration of training, and supervision. For more than half of the included studies however, the sample size of the sessions included to measure the inter-rater agreement leads to a high or unclear RoB.

Coding instruments

Below, we will review the seven included instruments on the ingredients they measure, how they measure these ingredients, and their psychometrics.

Global Rating of Motivational Interviewing Therapist (GROMIT)

The GROMIT [24] mainly concentrates on the relational ingredients, with an emphasis on MI-Spirit. It rates the therapist skill through fifteen 7-point scale-items, such as "The therapist directed the client's attention toward their own strengths". The extremes and the middle of the 7-point scales are defined: "Do not agree", "Somewhat agree", "Fully agree". The inter-rater agreement of the GROMIT-scales is fair to excellent [25] (Table 2).

Independent Tape Rating Scale (ITRS)

The ITRS [26], consists of 39 items to be scored on a 7-point scale, and addresses several potential active ingredients. Its main focus is on the technical ingredients, measuring MI-consistent (MICO) and MI-inconsistent (MIIN) therapist behavior. Two items measure relational ingredients, and are directed at MI-Spirit and at empathic understanding. Finally, two items evaluate the clients' motivational level. There are also IRTS-items that do not assess MI, but assess general substance abuse counseling interventions, and general characteristics of the therapists and the clients [27].

The twenty MI-items are scored on both adherence (1 = not at all, 7 = extensively) and competence (1 = very poor, 7 = excellent), leading to 42 scores including the two motivational level scores. The inter-rater agreement of those items is fair to excellent [26] (Table 3).

Motivational Interviewing Process Code (MIPC)

The MIPC is a training tool that consists of two lists ("functional skills", "dysfunctional skills") [28]. Both lists combine items directed at technical and at relational ingredients. Each item must be scored at a 5-point scale, for which all points are defined. The authors computed the percentage of inter-rater agreement, and they found low percentage of agreement [28] (Table 2).

Motivational Interviewing Skill Code 2.0 and 2.1 (MISC)

The MISC measures both therapist and client behavior [29,30]. It measures the relational ingredients by 7-point Global Counselor Rating scales, evaluating the extent to which the therapist communicates acceptance, empathy and MI-Spirit, and by one Global Client Rating ("client self-exploration").

For the technical ingredients, the coder counts the utterances of the therapist [29], and classifies these utterances in 19 categories that are either MICO (e.g. "reflect"), MIIN (e.g. "confront"), or neutral (e.g. "structure"). This also enables the coder to determine the therapist proficiency and the degree of intervention fidelity of the therapist, by calculating the summary scores of the MISC (e.g. percentage MICO responses).

Furthermore, the coder counts client responses (e.g. "expressing ability"), and determines the direction (towards or away from behavior change) of the change talk. For this, the MISC 2.0/2.1 incorporated the Commitment Language Coding System developed by Amrhein et al. [13,29,30]. All responses are categorized in eight codes for client behavior counts, reflecting the degree of the client's willingness, ability and readiness to change. The coding of the strength of client utterances is optional because in the MISC 2.0-version the reliability on these strength ratings was hard to establish [29].

Though the inter-rater agreement for the global ratings varies between studies [15,16,43], the high-quality study of Gaume et al. [16], showed that the training in scoring of the global ratings may lead to a fair to good level of agreement. Other studies [15,43] found mainly poor agreement on these global ratings (Table 2).

In the MISC, behavior counts of therapist and client show the total number of codes that each coder has assigned to specific behavior categories. The inter-rater agreement of the separate behavior counts of the subcategories of MICO and MIIN showed a pattern of wide variation between studies. Again, the studies of Gaume et al. [16,42] showed that training in the coding of separate MICO subcategories led to fair to excellent inter-rater agreement, while the coding of the MIIN behaviors are much more difficult to train, probably because of the rare occurrence of MIIN behavior in the coded sessions (poor to excellent) [16,42] (Table 2).

The eight client behavior codes are either change talk, sustain talk, or neutral. On these counts the ICC varied from good to excellent [42,43]. The inter-rater agreement for the average of the strength ratings of client behavior varied from poor to excellent [15,16] (Table 2).

Motivational Interviewing Supervision and Training Scale (MISTS)

The MISTS is primarily a training tool [32]. The instrument measures both relational and technical ingredients. The technical ingredients are measured through eight categories of therapist behavioral counts (e.g. "simple reflection"), and by some of the sixteen global ratings. The other global ratings measure relational ingredients (e.g. "collaborating with client"), focusing on therapist behavior. One global rating is directed towards client behavior, and one rates the fidelity to MI..

The global ratings are scored on 7-point scales, with defined anchors on point 1, 4, and 7. The inter-rater agreement for the global ratings is fair to excellent [32] (Table 2). There is no information on the inter-rater agreement on the behavioral counts.

Motivational Interviewing Treatment Integrity (MITI)

In the introduction to MITI 3.1.1 [33] the authors underline that the MITI is designed as a treatment integrity and feedback instrument. The MISC, the "parent instrument" of the MITI, is more useful for detailed MI-process research [33].

The recent 3.1.1-version of the MITI uses a random 20-minute section for coding and for scoring the global ratings. It measures relational ingredients by 5-point global rating scales. All anchors are defined on each scale. Technical ingredients are measured by therapist behavior counts, divided in eight categories. These categories focus on the most important therapist behaviors in MI. All above-mentioned measures contribute to calculate the summary scores, which reveal the proficiency and the fidelity in MI of the therapist.

Two studies [44,45] evaluated the inter-rater reliability of the MITI 3.1.1. In one study, the ICC for the global ratings is 0.20 [45]. This poor inter-rater reliability is probably influenced by the limited variability in the scores on the global ratings. Kaplan et al. [44] found substantial inter-rater agreement for all global ratings, in this study the measures were recoded as a match if the difference between the raters was one point on the 5-point scale. The inter-rater agreement scores for the therapist behavior counts are all excellent in one study [45], and poor to excellent in the second study [44]. These differences may be influenced by differences in coder training time (Table 2).

Motivational Interviewing Sequential Code for Observing Process Exchanges (SCOPE)

The SCOPE was developed to code and investigate sequential information on MI [37]. The SCOPE elaborates on the MISC, and adds the coding of direction (positive, neutral, or negative) to the questions and reflections of the therapist [37]. The SCOPE measures technical ingredients in context: the impact of the therapist behavior on the client, and vice versa, is visible through the sequential coding. The coder uses 19 therapist behavior codes, and nine client behavior codes. It is also possible to compute the same summary scores as in the MISC, to detect the MI-proficiency and the fidelity of the therapist.

Three studies [46,47,50] have described the psychometric properties of the SCOPE. Two of these studies [46,50] used the same sample, so the reliability of the SCOPE is computed in two studies [46,47]. These studies computed the reliability of the SCOPE at utterance-to-utterance level. A sequential coding system is reliable only if different coders assign the same code to the same utterance, whereas the reliability of the MISC, in which the codes usually are counted, refers to the agreement on the total score at session level. In the first small study [46], a moderate to good inter-rater agreement was found (Table 2). For the second study [47] the authors reported an average Kappa of 0.75 with a range of 0.56-0.87 on the behavior categories [18] (Table 2).

Table 3 offers an oversight of the potential of the instruments to measure relevant information of MI-sessions.

Discussion

Our review suggests that a combination of instruments reliably measures different potential active ingredients of the interactive MI-process. According to Miller & Rollnick [7,8] the three fundamental characteristics of MI are: "(1) a person-centered, non-authorian counseling style (…), and (2) a clearly identified change goal (…), and (3) differential evoking and strengthening of the person's own motivation for change." [8, p235]. This suggests that the relational active ingredients contribute to the first characteristic, and that the instrument should measure MI-Spirit, empathic communication, and client-therapist relationship. The technical active ingredients should contribute to the third characteristic, therefore, the instrument should measure the techniques and strategies to evoke and strengthen the client change talk and diminish sustain talk. The instrument should also measure the effect of these techniques and strategies on the client: does the change talk increase and become stronger?

Please cite this article as: Dobber J, et al, Selecting an optimal instrument to identify active ingredients of the motivational interviewing-process, J Psychosom Res (2014), http://dx.doi.org/10.1016/j.jpsychores.2014.10.010

Table 1
Instruments

Instrument	Inclusion criteria*				Short description of the instrument	Goal of the instrument
	1	2	3	4		
BECCI (Behavior Change Counseling Index) Lane et al. [19]	−	+	+	+	Eleven 5-point scales on therapist skills.	To measure practitioner competence in behavior change counseling (BCC), an adaptation of motivational interviewing suitable for brief consultations in healthcare settings.
CASPI (Computer Assessment of Simulated Patient Interviews) Baer et al. [20]	+	−	+	+	A combination of dichotomous codes and a 5-point scale on therapist skills: reflective listening, responding to sustain talk, responding to change talk, eliciting change talk, affirming, summarizing.	To assess MI skills through a web-based assessment.
CBCCAI (Combined Behavioral Change Counseling Assessment Instrument) Strayer et al. [21]	+	+	+	−	Twenty-three closed (yes/no) questions on components, therapist tasks and therapist skills.	To evaluate the fidelity and quality of brief behavioral change interventions based on the 5A's, Stages of Change, or MI.
CEMI (Client Evaluation of Motivational Interviewing) Madson et al. [22]	+	+	−	+	Thirty-five 4-point scales on therapist behavior, rated by the client.	To provide feedback and basis for supervision by assessing client perception of clinician MI use.
CLEAR (Client Language Easy Rating Coding System) Glynn & Moyers [23]	+	+	+	−	Tallies of client behavior, divided in change talk and counterchange talk.	To classify and quantify client language that is either change talk or counter-change talk.
GROMIT (Global Rating of Motivational Interviewing Therapist) Moyers [24]; Resko et al. [25]	+	+	+	+	Fifteen 7-point scales on therapist skill en MI-competence.	To measure MI-therapist skill, responsiveness and overall competence.
ITRS (Independent Tape Rating Scale) Martino et al. [26,27]	+	+	+	+	Thirty-nine 7-point scales, thirty-seven on therapist adherence and competence to MI or common drug counseling, and general therapist and two on client motivation. The 37 therapist-items are scored twice: on adherence and on competence.	To evaluate the therapists use of MI strategies, techniques inconsistent with MI, and general substance abuse monitoring.
MIPC (Motivational Interviewing Process Code) Barsky & Coleman [28]	+	+	+	+	Thirteen 5-point scales on functional MI-skills, and twelve 5-point scales on dysfunctional MI-skills.	To measure student competencies in MI skills.
MISC 2.0/2.1 (Motivational Interviewing Skill Code) Miller et al. [29,30]	+	+	+	+	Three 7-point scales to score the global impression of the therapist on Acceptance, Empathy, and MI-Spirit; One 7-point scale to score the client self-exploration; Behavior counts on therapist and client utterances; Strength coding of client utterances; Coding of direction of client utterances (towards or away from the target behavior); Summary scores, indicating the quality of MI.	To evaluate the quality of MI from audiotapes and videotapes of individual counseling sessions.

MIST-ED (Motivational Interviewing Scenarios Tool for Eating Disorders Sepulveda et al. [31]	+	−	+	+	Nine response categories (4 MI adherent, 4 MI non-adherent, 1 other) to classify the statements.	To assess the MI-skills of caregivers of adolescents with eating disorders.
MISTS (Motivational Interviewing Supervision and Training Scale) Madson et al. [32]	+	+	+	+	Behavior counts of types of therapist responses uttered during sessions; Sixteen 7-point scales on the quality, MI fidelity and effectiveness on therapist interventions.	To assist in training and supervision of therapists by measuring the quality, fidelity and effectiveness of the MI sessions.
MITI 3.1.1 (Motivational Interviewing Treatment Integrity) Moyers et al. [33]	+	+	+	+	Three 5-point scales to capture the overall impression on MI-Spirit (a composition of 3 sub-scales), Empathy and Direction. Behavior counts of therapist utterances divided in Giving Information, MI Adherent, MI Non-adherent, Question (open/closed), Reflection (simple/complex). Therapist proficiency summary scores.	To evaluate the competence of the therapist in performing MI.
PCCCS (Patient-Centered Communication Coding System) Ledoux et al. [34]	+	+	+	+	Four 5-point scales to capture the overall impression on Collaboration, Autonomy, Direction, and Empathy; Behavior counts on 12 categories on (positive/negative) therapist utterances, based on Patient-Centered Communication.	To assess patient-centered communication techniques as a process evaluation of fidelity.
PEPA (Peer Proficiency Assessment) Mastroleo et al. [35]	+	+	+	−	Behavior counts of Questions (open/closed) and Reflections (simple/complex).	To examine MI-adherence in undergraduate student peer delivered interventions.
REM (Rating Scales for the Assessment of Empathetic Communication in Medical Interviews) Nicolai et al. [36]	+	+	−	+	Nine 7-point scales, rated by the client. Six of these nine scales are directed to Empathy, three scales are directed to Confrontation.	To assess empathy and confrontation in physician-patient interactions.
SCOPE (Motivational Interviewing Sequential Code for Observing Process Exchanges) Martin et al. [37]	+	+	+	+	Sequential coding of therapist utterances and client utterances; Coding of direction of therapist and client utterances (towards or away from the target behavior); Summary scores, indicating the quality of MI.	To encode recorded and transcripted MI interactions between a therapist and an individual client, with a particular focus on the sequential information contained in the exchange between the parties, for the purpose of investigating the relationship between theoretical constructs important to MI, therapy process more generally, and client outcome.
VASE/VASE-R (Video Assessment of Simulated Encounters) Rosengren et al. [38,39]	+	−	+	+	Five subscales on Reflective Listening, Responding to Resistance, Summarizing, Eliciting Change Talk, Developing Discrepancy. Together the five subscales comprise eighteen 3-point items.	To assess the overall MI skill and 5 MI microskills through video vignettes.

* 1. The instrument specifically addresses measuring the execution of MI; 2. the instrument brings into focus one or more active ingredients in the MI-process, and/or its effect on client behavior; 3. the measurements are based on observations; 4. the instrument collects detailed information.

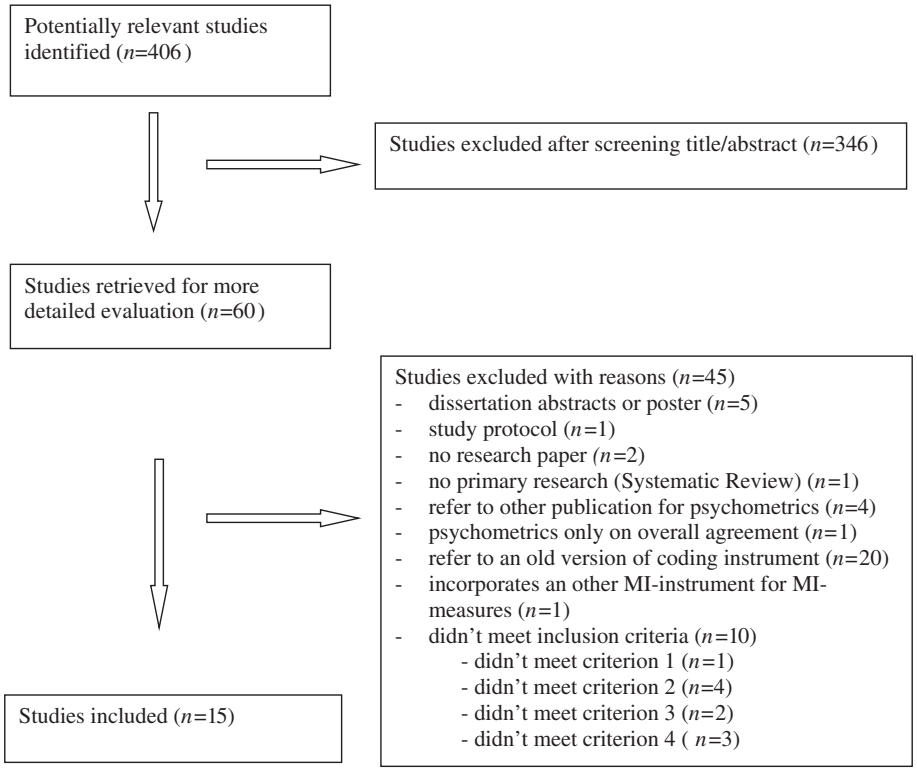

Fig. 1. Flow diagram.

All instruments, except the SCOPE, measure potential relational active ingredients. These ingredients are measured by global rating scales, which are directed at MI-Spirit and at empathic understanding, although not always explicitly. None of the instruments has a direct measurement of the client-therapist relationship. And all instruments, except the GROMIT, measure potential technical active ingredients. Two instruments use global rating scales [26,28], three instruments measure the technical ingredients by behavior counts [29,30,32,33], and one instrument uses sequential coding [37]. Rating scales can only give an impression of the overall use of techniques (e.g. "Open ended questions" [26]), and, therefore, they don't show the interactive process of the MI-session. Counts of therapist behavior provide insight into intervention fidelity and therapist proficiency. If the counts are linked to client behavior counts, they may show associations between the use of certain techniques and the proportion of client change talk and sustain talk. However, this doesn't reveal the immediate effect of therapist behavior on client behavior. For detailed process information, it is best that the order in which the behaviors of the therapist and the client occur has been retained. The sequential coding of the SCOPE provides this detailed information on therapist behaviors, on the impact on the client, and on adaptations of the therapists' strategies based on the client reactions. Also, the SCOPE reveals the direction of the questions and reflections of the therapist, which may facilitate interpretations on successful therapist strategies. The MISC can also be used for sequential coding, but the authors of the MISC advise the use of the SCOPE for sequential coding, as an instrument that reveals detailed information on the therapy process [29]. The MISC is the only instrument that measures the strength of client statements. Although it is hard to measure these strength ratings in a reliable way, several studies found associations between the strength of statements and client outcomes [13,15,16], so strength rating adds extra detail to the measurement of client behavior.

In their paper on the impact of treatment fidelity on the (in)effectiveness of complex behavioral interventions, Miller & Rollnick [8] stress the importance of the deliverance of the right intervention content. Therapist proficiency and intervention fidelity will probably enlarge the presence of active ingredients in the intervention, and therefore, though the quality of the MI delivered is not an active ingredient in itself, the measurement of the fidelity may help to interpret the research findings. Five instruments measure this intervention fidelity, by global rating scales [26,32] or, more detailed, by summary scores [29,33,37].

The active ingredients must be valid and reliably measured. The inter-rater reliability of three instruments is within acceptable range and is computed under circumstances with low risk of bias (Table 2). Of these, the GROMIT and the MISC both measure potential relational active ingredients, but 14 of the 16 global ratings of the GROMIT reach a good inter-rater agreement [25], while the inter-rater agreement of the four global ratings of the MISC is mostly fair [16]. The MISC and the SCOPE measure potential technical active ingredients. The reliability of the SCOPE on an utterance-to-utterance level is moderate to good. However, the research on the selected instruments is scarce. We found no studies establishing the validity of the GROMIT, MISC 2, and SCOPE. The studies we have found only concentrated on the reliability of the instruments. For the GROMIT, we found only one RCT in which the inter-rater agreement [25] of the GROMIT was established. For the MISC, the psychometrics on the MISC 2-versions rely heavily on the research by one research group [16,40–42]. The studies on the SCOPE show moderate to good reliability, but this instrument has only been tested by its developers. In addition, it must be taken in account that reliably assessing the strength of client speech is difficult, and there is a wide range of the *average* strength ratings (ICC-range 0.38 to 0.78) [15,16]. We could not find information on the reliability of strength coding on an utterance level. Finally, for statistical reasons, most studies have categorized the separate behaviors, but differ in the composition of these categories. These are limitations of the present state of the art, and, although it hinders the interpretation and the comparison of the psychometrics between different studies, most of the values of the psychometrics are in the same range. This means that they are trustworthy enough to rely on for a decision on the choice of a research instrument.

Please cite this article as: Dobber J, et al, Selecting an optimal instrument to identify active ingredients of the motivational interviewing-process, J Psychosom Res (2014), http://dx.doi.org/10.1016/j.jpsychores.2014.10.010

Table 2
Risk of bias[1] and inter-rater agreement

Study	Instrument	Size reliability sample[2]	Sampling method[3]	Duration of coder training[4]	Supervision[5]	Inter-rater agreement (ICC-range[6])
Moyers [24] and Resko et al. [25]	GROMIT	L	L	L	L	Global ratings: .41–.82
Martino et al. [26]	ITRS	H	L	L	U	MI-consistent skills: .66–.99 MI-inconsistent skills: .55–.98 Client motivation: .96–.96
Barsky & Coleman [28]	MIPC	U	H	U	U	percentage[7] inter-rater agreement functional skills: 51.27% percentage inter-rater agreement dysfunctional skills: 75.03%
Gaume et al. [16,40] and Bertholet et al. [41]	MISC 2.0	L	L	L	L	Global ratings: .50–.62 Therapist MI-consistent: .56–.82 Therapist MI-inconsistent: .22–.48 Therapist neutral: .36–.83 Client behavior: .71–.77 Strength of change talk[8]: 38–.75
Campbell et al. [15]	MISC 2.0 modified	H	H	H	L	Global ratings: poor[9] Client behavior: .75–.80 Strength of change talk[10]: .50–.78
Gaume et al. [42]	MISC 2.1	L	L	L	L	Therapist MI-consistent: .43–.91 Therapist MI-inconsistent (total): .79 Therapist neutral: .70–.89 Client behavior: .66–.79
Vader et al. [43]	MISC 2.1	L	L	L	U	Global ratings: –.20–.67 Therapist MI-consistent (total): .96 Therapist MI-inconsistent (total): .07 Client behavior: .84–.87
Kaplan et al. [44]	MISC 2.1	U	U	U	L	Client behavior: .72–.74
Madson et al. [32]	MISTS	L	L	H	U	Specific active listening skills: .41–.81 Specific skills MI-Spirit: .45–.74 Overall therapist ratings: .66–.76
Seng & Lovejoy [45]	MITI 3.1.1	H	U	L	L	Global ratings: .20 Behavior counts: .77–.90
Kaplan et al. [44]	MITI 3.1.1	U	U	H	L	Global ratings: .61–.74[11,12] Behavior counts: .18–.76[6]
Moyers & Martin [46]	SCOPE	H	L	L	L	Therapist MI-consistent: .66[11] Therapist MI-inconsistent: .68[11] Therapist – other behavior: .82[11] Client change talk: .79[11] Client counter change talk: .60[11] Client – neutral/ask: .79[11]
Moyers et al. [47]	SCOPE	L	L	L	U	Sequential coding of utterances: .56–.87[11] Frequency therapist MI-consistent: .49–.98[6] Frequency therapist MI-inconsistent: .79[6] Frequency client behavior: .88–.96[6]

[1] Risk of Bias: L = low risk of bias; H = high risk of bias; U = Uncertain risk of bias.
[2] Proportion sample size: L = a proportion of at least 20%; H = a proportion of less than 20%; U = proportion not reported.
[3] Sampling method: L = all sessions or random; H = non-random methods; U = sampling method not reported.
[4] Duration of coder training: L = 35 h of more, or training until sufficient inter-rater agreement was achieved; H = <35 h; U = duration of coder training not reported.
[5] Supervision: L = supervision or coder meetings; H = no supervision or coder meetings; U = supervision or coder meetings not reported.
[6] ICC = Intraclass Correlation Coefficient. The interpretation of the ICC is: below 0.40 = poor; 0.40–0.59 = fair; 0.60–0.74 = good; 0.75–1.00 = excellent [48].
[7] ICC not computed.
[8] Average change talk strength (+5 to −5).
[9] All Global ratings were poor. ICCs not reported.
[10] Average change talk strength (+3 to −3).
[11] Kappa, not ICC. The interpretation of Kappa is: below 0.21 = poor; 0.21 to 0.40 = fair; 0.41–0.60 = moderate; 0.61 to 0.80 = substantial; 0.81–1.00 = good [49].
[12] Measures were recoded as a match if the measure between raters differed by one increment on this 5-point scale.

Table 3
Summary table

Instrument	Potential relational active ingredients	Potential technical active ingredients	Client behavior	Sequential coding	Strength coding	MI-quality/proficiency/fidelity
GROMIT [24]	X					
ITRS [26,27]	X	X				X
MIPC [28]	X	X				
MISC 2.0/2.1 [29,30]	X	X	X		X	X
MISTS [32]	X	X				X
MITI 3.1.1 [33]	X	X				X
SCOPE [37]		X	X	X		X

Please cite this article as: Dobber J, et al, Selecting an optimal instrument to identify active ingredients of the motivational interviewing-process, J Psychosom Res (2014), http://dx.doi.org/10.1016/j.jpsychores.2014.10.010

Conclusions

In conclusion, the potential relational active ingredients can best be measured by the global ratings of the GROMIT, with global ratings of the MISC as an alternative option. The potential technical active ingredients can best be measured by the SCOPE, or by the MISC. These two instruments measure both therapist behavior and client behavior. The method of behavior counts though, employed by the MISC, offers less information than the sequential coding of the SCOPE. The SCOPE also makes the direction of therapist behavior visible, while strength ratings of client speech is only measured by the MISC. For the quality measuring of the intervention delivered, the summary scores of the MISC can be used.

We propose that future research applies a comprehensive approach to link the SCOPE as the only instrument for sequential coding, and the global ratings of GROMIT or of the MISC, the strength ratings of the MISC 2.1, and the summary scores of the MISC, to client outcomes. This can be used to evaluate the effectiveness of techniques, client-therapist relationship and empathic communication, which will lead to more effective use of MI, which in turn may lead to better outcomes for clients in clinical practice.

Conflict of interest statement

All co-authors have approved the publication of the manuscript, and have no competing interests to report.

Acknowledgements

We would like to thank Lori Broere MSc for her help in the searching, selecting, quality assessment and data extraction, and Marjolein Snaterse MSc and Vera Kruse MSc for their help in the quality assessment and data-extraction.

References

[1] Hettema J, Steele J, Miller WR. Motivational interviewing. Annu Rev Clin Psychol 2005;1:91–111.
[2] Rubak S, Sandbaek A, Lauritzen T, Christensen B. Motivational interviewing: A systematic review and meta-analysis. Br J Gen Pract 2005;55:305–12.
[3] Burke BL, Arkowitz H, Menchola M. The efficacy of motivational interviewing: A meta-analysis of controlled clinical trials. J Consult Clin Psychol 2003;71:843–61.
[4] Hunt GE, Siegfried N, Morley K, Sitharthan T, Cleary M. Psychosocial interventions for people with both severe mental illness and substance misuse. Cochrane Database Syst Rev 2013. http://dx.doi.org/10.1002/14651858.CD001088.pub3 [Art.No.:CD001088].
[5] Lundahl B, Burke BL. The effectiveness and applicability of motivational interviewing: A practice-friendly review of four meta-analyses. J Clin Psychol In Session 2009;65:1232–45.
[6] Barrowclough C, Haddock G, Tarrier N, Lewis SW, Moring J, O'Brien R, et al. Randomized controlled trial of motivational interviewing, cognitive behavior therapy, and family intervention for patients with comorbid schizophrenia and substance use disorders. Am J Psychiatry 2001;158:1706–13.
[7] Miller WR, Rollnick S. Motivational interviewing. Helping people change3rd ed. . New York: Guilford Press; 2013.
[8] Miller WR, Rollnick S. The effectiveness and ineffectiveness of complex behavioral interventions: impact of treatment fidelity. Contemp Clin Trials 2014;37:234–41.
[9] Morgenstern J, Kuerbis A, Amrhein P, Hail L, Lynch K, McKay JR. Motivational interviewing: A pilot test of active ingredients and mechanisms of change. Psychol Addict Behav 2012;26:859–69.
[10] Apodaca TR, Longabaugh R. Mechanisms of change in motivational interviewing: A review and preliminary evaluation of the evidence. Addiction 2009;104:705–15.
[11] Miller WR, Rose GS. Toward a theory of motivational interviewing. Am Psychol 2009;64:527–37.
[12] Miller WR, Benefield GR, Tonigan JS. Enhancing motivation for change in problem drinking: A controlled comparison of two therapists styles. J Consult Clin Psychol 1993;61:455–61.
[13] Amrhein PC, Miller WR, Yahne CE, Palmer M, Fulcher L. Client commitment language during motivational interviewing predicts drug outcomes. J Consult Clin Psychol 2003;71:862–78.
[14] Baer JS, Beadnell B, Garrett SB, Hartzler B, Wells EA, Peterson PL. Adolescent change language within a brief motivational intervention and substance use outcomes. Psychol Addict Behav 2008;22:570–5.
[15] Campbell SD, Adamson SJ, Carter JD. Client language during motivational enhancement therapy and alcohol use outcome. Behav Cogn Psychother 2010;38:399–415.
[16] Gaume J, Gmel G, Daeppen JB. Brief alcohol interventions: Do counsellors' and patients' communication characteristics predict change? Alcohol Alcohol 2008;43:62–9.
[17] Martin T, Christopher PJ, Houck JM, Moyers TB. The structure of client language and drinking outcomes in Project MATCH. Psychol Addict Behav 2011;25:439–45.
[18] Madson MB, Campbell TC. Measures of fidelity in motivational enhancement: A systematic review. J Subst Abuse Treat 2006;31:67–73.
[19] Lane C, Huws-Thomas M, Hood K, Rollnick S, Edwards K, Robling M. Measuring adaptations of motivational interviewing: The development and validation of the behavior change counseling index (BECCI). Patient Educ Couns 2005;56:166–73.
[20] Baer JS, Carpenter KM, Beadnell B, Stoner SA, Ingalsbe MH, Hartzler B, et al. Computer Assessment of Simulated patient Interviews (CASPI): Psychometric properties of a web-based system for the assessment of motivational interviewing skills. J Stud Alcohol Drugs 2012;73:154–64.
[21] Strayer SM, Martindale JR, Pelletier SL, Rais S, Powell J, Schorling JB. Development and evaluation of an instrument for assessing brief behavioral change interventions. Patient Educ Couns 2011;83:99–105.
[22] Madson MB, Mohn RS, Zuckoff A, Schumacher JA, Kogan J, Hutchison S, et al. Measuring client perceptions of motivational interviewing: factor analysis of the Client Evaluation of Motivational Interviewing scale. J Subst Abuse Treat 2013;44:330–5.
[23] Glynn LH, Moyers TB. [Internet] Manual for the Client Language EAsy Rating Coding System: Formerly: "Motivational Interviewing Skill Code (MISC) 1.1". [cited 2014 January 9] Available from http://casaa.unm.edu/download/CLEAR.pdf; 2012.
[24] Moyers TB. [Internet] Global Rating of Motivational Interviewing Therapist (GROMIT). [cited 2013 March 28] Available from http://casaa.unm.edu/download/gromit.pdf; 2004.
[25] Resko SM, Walton MA, Chermack ST, Blow FC, Cunningham RM. Therapist competence and treatment adherence for a brief intervention addressing alcohol and violence among adolescents. J Subst Abuse Treat 2012;42:429–37.
[26] Martino S, Ball SA, Nich C, Frankforter TL, Carroll KM. Community program therapist adherence and competence in motivational enhancement therapy. Drug Alcohol Depend 2008;96:37–48.
[27] Martino S, Ball S, Nich C, Frankforter TL, Carroll KM. Correspondence of Motivational Enhancement Treatment Integrity Ratings among therapists, supervisors, and observers. Psychother Res 2009;19:181–93.
[28] 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.
[29] Miller WR, Moyers TB, Ernst D, Amrhein P. [Internet] Manual for the Motivational Skill Code (MISC). Version 2.1. [cited 2013 October 27] Available from http://casaa.unm.edu/download/misc.pdf; 2008.
[30] Miller WR, Moyers TB, Ernst D, Amrhein P. [Internet] Manual for the Motivational Skill Code (MISC). Version 2.0. [cited 2013 March 28] Available from http://casaa.unm.edu/download/misc2.pdf; 2003.
[31] Sepulveda AR, Wise C, Zabala M, Todd G, Treasure J. Development and reliability of a Motivational Interviewing Scenarios Tool for Eating Disorders (MIST-ED) using a skills-based intervention among caregivers. Eat Behav 2013;14:432–6.
[32] Madson MB, Campbell TC, Barrett DE, Brondino MJ, Melchert TP. Development of the Motivational Interviewing Supervision and Training Scale. Psychol Addict Behav 2005;19:303–10.
[33] Moyers TB, Martin T, Manuel JK, Miller WR, Ernst D. [Internet] Motivational Interviewing Treatment Integrity 3.1.1 (MITI 3.1.1). [cited 2013 March 28] Available from http://casaa.unm.edu/download/MITI3_1.pdf; 2010.
[34] Ledoux T, Hilmers A, Watson K, Baranowski T, O'Connor TM. Development and feasibility of an objective measure of patient-centered communication fidelity in a pediatric obesity intervention. J Nutr Educ Behav 2013;45:349–54.
[35] Mastroleo NR, Mallet KA, Turrisi R, Ray AE. Psychometric properties of the Peer Proficiency Assessment (PEPA): A tool for evaluation of undergraduate peer counselors' motivational interviewing fidelity. Addict Behav 2009;34:717–22.
[36] Nicolai J, Demmel R, Hagen J. Rating Scales for the assessment of Empathic communication in Medical interviews (REM): Scale development, reliability, and validity. J Clin Psychol Med Settings 2007;14:367–75.
[37] Martin T, Moyers TB, Houck J, Christopher P, Miller WR. [Internet] Motivational Interviewing Sequential Code for Observing Process Exchanges (SCOPE). Coder's Manual. [cited 2013 march 28] Available from http://casaa.unm.edu/download/scope.pdf; 2005.
[38] Rosengren DB, Baer JS, Hartzler B, Dunn CW, Wells EA. The video assessment of simulated encounters (VASE): Development and validation of a group-administered method for evaluating clinician skills in motivational interviewing. Drug Alcohol Depend 2005;79:321–30.
[39] Rosengren DB, Hartzler B, Baer JS, Wells EA, Dunn CW. The video assessment of simulated encounters-revised (VASE-R): Reliability and validity of a revised measure of motivational interviewing skills. Drug Alcohol Depend 2008;97:130–8.
[40] Gaume J, Gmel G, Faouzi M, Daeppen JB. Counsellor behaviours and patient language during brief motivational interventions: A sequential analysis of speech. Addiction 2008;103:1793–800.
[41] Bertholet N, Faouzi M, Gmel G, Gaume J, Daeppen JB. Change talk sequence during brief motivational intervention, towards or away from drinking. Addiction 2010;105:2106–12.

Please cite this article as: Dobber J, et al, Selecting an optimal instrument to identify active ingredients of the motivational interviewing-process, J Psychosom Res (2014), http://dx.doi.org/10.1016/j.jpsychores.2014.10.010

[42] Gaume J, Bertholet N, Faouzi M, Gmel G, Daeppen JB. Counselor motivational interviewing skills and young adult change talk articulation during brief motivational interventions. J Subst Abuse Treat 2010;39:272–81.
[43] Vader AM, Walters ST, Prabhu GC, Houck JM, Field CA. The language of motivational interviewing and feedback: Counselor language, client language, and client drinking outcomes. Psychol Addict Behav 2010;24:190–7.
[44] Kaplan JE, Keeley RD, Engel M, Emsermann C, Brody D. Aspects of patient and clinician language predict adherence to antidepressant medication. J Am Board Fam Med 2013;26:409–20.
[45] Seng EK, Lovejoy TI. Reliability and validity on a treatment fidelity assessment for motivational interviewing targeting sexual risk behaviors in people living with HIV/AIDS. J Clin Psychol Med Settings 2013;20:440–8.
[46] Moyers TB, Martin T. Therapist influence on client language during motivational interviewing sessions. J Subst Abuse Treat 2006;30:245–51.
[47] Moyers TB, Martin T, Houck JM, Christopher PJ, Tonigan JS. From in-session behaviors to drinking outcomes: A causal chain for motivational interviewing. J Consult Clin Psychol 2009;77:1113–24.
[48] Cicchetti DV. Guidelines, criteria and rules of thumb for evaluating normed and standardized instruments in psychology. Psychol Assess 1994;6:284–90.
[49] Petrie A, Sabin C. Medical statistics at a glance. 2nd ed. Oxford: Blackwell Publishing; 2005.
[50] Moyers TB, Martin T, Christopher PJ, Houck JM, Tonigan JS, Amrhein P. Client language as a mediator of motivational interviewing efficacy: Where is the evidence? Alcohol Clin Exp Res 2007;31:40S–7S.

Please cite this article as: Dobber J, et al, Selecting an optimal instrument to identify active ingredients of the motivational interviewing-process, J Psychosom Res (2014), http://dx.doi.org/10.1016/j.jpsychores.2014.10.010