

CONTENT VALIDITY AND COGNITIVE TESTING IN THE DEVELOPMENT OF A MOTIVATIONAL INTERVIEWING SELF-ASSESSMENT QUESTIONNAIRE

VSEBINSKA VELJAVNOST IN KOGNITIVNO TESTIRANJE PRI RAZVOJU SAMO-OCENJEVALNEGA VPRAŠALNIKA O MOTIVACIJSKEM INTERVJUJU

Tadeja HOČEVAR^{1*}, Tim ANTISS², Danica ROTAR PAVLIČ³

¹ National Institute of Public Health, Trubarjeva 2, 1000 Ljubljana, Slovenia

² University of Reading, Henley Business School, Whiteknights Rd, Reading RG6 6UD, United Kingdom

³ University of Ljubljana, Medical Faculty, Department of Family Medicine, Poljanski nasip 58, 1000 Ljubljana, Slovenia

Received: Oct 05, 2023
Accepted: Nov 28, 2023

Original scientific article

ABSTRACT

Keywords:

Alcohol screening and brief intervention
Cognitive testing
Content validity
Motivational interviewing
Self-assessment questionnaire

Aim: To develop and content validate a self-assessment questionnaire on motivational interviewing (MI) practice as the first stages in forming the questionnaire to be used in cross-sectional studies involving practitioners conducting the MI-based alcohol screening and brief intervention (ASBI).

Methods: A comprehensive mixed methods approach included a literature review, 3 rounds of expert panel (EP) opinions (n=10), cognitive testing (CT) with 10 MI-based ASBI practitioners, and questionnaire piloting with 31 MI-based ASBI practitioners. Based on the EP opinions in the second round, content validity indices (CVIs) and the modified kappa coefficient (k*) were calculated, focusing on the relevance and understandability of questions and comprehensiveness and meaningfulness of the response options. This analysis was performed in 2020, at the conclusion of the national "Together for a Responsible Attitude Towards Alcohol Consumption" ("Skupaj za odgovoren odnos do pitja alkohola", SOPA) project's pilot implementation.

Results: On a scale level, CVI values based on universal agreement for the entire questionnaire were high for 3/4 categories (S-CVI-UA>0.80), and CVI values based on average agreement were high across all categories (S-CVI-Ave>0.90). At the item level, CVI values (I-CVI) were never <0.50 (automatic item rejection), and the modified kappa value (k*) indicated poor validity for two items in the understandability category (k*=0.33). All problematic parts of the questionnaire were further tested and successfully modified based on the results of CT, and accepted in the third round of testing.

Conclusions: The final version of the questionnaire demonstrated appropriate content validity for use in studies among Slovenian MI-based ASBI practitioners and is now ready for further psychometric testing.

IZVLEČEK

Ključne besede:

kratko svetovanje za opuščanje tvegane in škodljivega pitja alkohola
kognitivno testiranje vsebinske veljavnosti motivacijskega intervju samo-ocenjevalni vprašalnik

Namen: Razviti samo-ocenjevalni vprašalnik o izvajanju motivacijskega intervjuja (MI) in preveriti njegovo vsebinsko veljavnost oz. izvesti prvi dve fazi oblikovanja vprašalnika za uporabo v presečnih raziskavah med izvajalci na MI temelječega kratkega svetovanja za opuščanje tvegane in škodljivega pitja alkohola (KSTŠPA).

Metode: Izvedli smo celovit pristop mešanih metod, ki je vključeval pregled literature, tri kroge mnenj skupine strokovnjakov s področja (SS) (n = 10), kognitivno testiranje (KT) vprašalnika z 10 izvajalci in pilotno testiranje vprašalnika s 31 izvajalci na MI temelječega KSTŠPA. Na podlagi mnenj SS v drugem krogu testiranja smo izračunali indekse vsebinske veljavnosti (IVV) in modificirani koeficient kapa (k*). Osredotočili smo se na 4 vsebinske kategorije: bistvenost in razumljivost vprašanj ter smiselnost in izčrpnost možnosti odgovorov. Razumljivost vprašalnika smo dodatno preverjali s kognitivnim testiranjem. Analizo smo izvedli v letu 2020 ob zaključku izvajanja pilota nacionalnega projekta Skupaj za odgovoren odnos do pitja alkohola (SOPA).

Rezultati: Vrednosti IVV na podlagi univerzalnega strinjanja strokovnjakov za celotni vprašalnik so bile ustrezno visoke (> 0,80) v 3 od 4 kategorij, IVV na podlagi povprečnega strinjanja strokovnjakov pa je bila visoka (> 0,90) v vseh kategorijah. Vrednosti IVV na ravni postavk niso bile manjše od 0,50, kar bi pomenilo avtomatično zavrnitev postavke, vrednosti k* pa so pokazale slabo veljavnost pri dveh postavkah v kategoriji razumljivost (k* = 0,33). Vse problematizirane dele vprašalnika smo dalje kognitivno testirali in na podlagi rezultatov uspešno spremenili ter so bili nato sprejeti kot ustrezni v tretjem krogu testiranja.

Zaključki: Končna različica vprašalnika ima ustrezno vsebinsko veljavnost za uporabo med slovenskimi izvajalci na MI temelječega KSTŠPA in je pripravljena na preverjanje psihometričnih lastnosti.

1 INTRODUCTION

Motivational interviewing (MI) is a widely used and effective conversational approach for helping people change their behaviour (1). It seeks to strengthen a person’s self-determined motivation by evoking their inner resources and strengths (1). Several studies indicate its effectiveness in counselling for excessive alcohol consumption. (2).

An increasing number of primary healthcare and other professionals in the helping professions use this approach, and assessing its quality or use is essential for programme and outcome optimisation. Tools for evaluating the integrity of MI practice include the MISC (Manual for the Motivational Interviewing Skill Code) (3-5) and the MITI (Motivational Interviewing Treatment Integrity Code) (6, 7). These tools involve an expert assessing (part of) the counselling session and the related demands with regard to time, financial sources and knowledge (6, 7). Another tool is supervisory, MIA:STEP (Motivational interviewing assessment: Supervisory tools for enhancing proficiency) (8), which can be self-administered by the practitioner and used for subsequent supervision and discussion. This also addresses a single session or part of a session. To the best of our knowledge, there are no comprehensive, easy to administer and validated self-assessment questionnaires regarding MI practice, which might be used to help self-assess longer time periods of MI usage to help shape practice and inform research on the effectiveness and outcomes in a practical way.

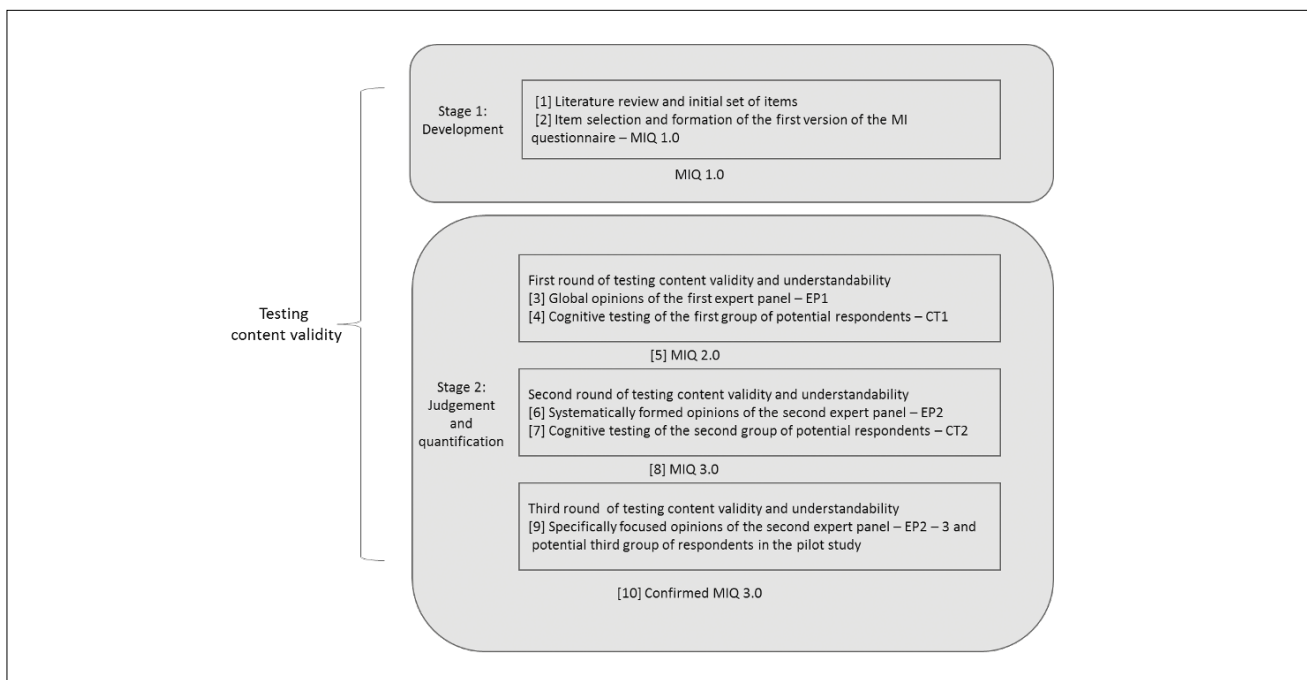
This study aims to address this gap through the initial development and content validation of a comprehensive self-assessment questionnaire to be used as an instrument in cross-sectional studies among Slovenian experts that use MI in their work (MI practitioners), focusing specifically on those conducting alcohol screening and brief intervention (ASBI) in primary healthcare settings and social work centres.

2 METHODS

A mixed methods approach was applied. We adhered to the content validity protocol as described by Lynn (9), upgraded with cognitive testing procedures. This iterative process involved 10 steps spanning two distinct stages (Figure 1).

2.1 Stage one - questionnaire development

The authors of this paper conducted a comprehensive literature review, including the foundational work of the MI authors Miller and Rollnick (1), literature on the main MI practice coding systems (3-8), and a literature review focusing on self-assessment of MI practice. This helped generate the initial pool of items for the first version of the questionnaire.



Legend: [] - numbers in brackets define steps in testing content validity;MIQ 1.0/2./3.0 - version of MI questionnaire

Figure 1. Stages and steps in testing the content validity of the MI practice self-assessment questionnaire.

The literature review was conducted using the PubMed bibliographic database in 2018, 2020, and during the summer of 2023. Keywords used in the title/abstract search included “self-evaluation questionnaire”, “self-assessment scale”, “self-evaluation”, and “self-assessment”. An article was considered relevant if it was an original research paper or a scientific review article that discussed self-assessment questionnaires related to the practice of MI. We excluded articles that focused on self-assessment of health outcomes in patients/clients or that were self-assessment questionnaires not specific to the practice of MI (e.g., attitudes toward practicing MI, its effects, satisfaction with MI training, etc.). We assessed the following data in the articles: the purpose of the tool, including who it was intended for and the time period it assessed, the MI elements selected, number of items, response categories defined, and number of response options on the response scales.

2.2 Stage two - judgment and quantification

2.2.1 Participants, materials, procedures, data collection, and analysis regarding expert panels

We established two expert panels for our study. The first panel consisted of five MI experts, four of whom were foreign members of the Motivational Interviewing Network of Trainers, one being the second author of this article. The fifth member was a national expert who had collaborated in the national project “Together for a Responsible Attitude Towards Drinking Alcohol” (TRATAC, and in Slovenian “Skupaj za odgovoren odnos do pitja alkohola”, SOPA) and helped to deliver MI-based ASBI training for primary healthcare and social workers. We sought the experts’ opinions on each item and the questionnaire as a whole, considering four perspectives: the relevance and understandability of the questions and comprehensiveness and meaningfulness of the response options. We used 4-point response scales (1 = not 2 = somewhat 3 = quite 4 = highly relevant/understandable/comprehensive/meaningful). The text for foreign MI experts was translated into English by a Slovenian-English translator and by the first author of this article, then proofread by the second author. Email was used to both interact with the experts and administer the questionnaire. The second round of the expert panel involved six national experts, all of whom were SOPA project MI trainers, with one having previously participated in the first round. Testing with both rounds of expert panels took place in the fall of 2020, with a three-week gap between rounds. In the first round, not all the experts completed the feedback form in its entirety; two of them provided more general opinions. Consequently, during data analysis, we considered general comments and removed items if at least one expert deemed them irrelevant. In the second round, all the experts completed the entire form and

provided more general opinions. Data analysis included calculation of three content validity indices, following the guidelines of Lynn (9), Polit (10), and Halek (11): the item-level content validity index (I-CVI) and both versions of the scale-level content validity indexes (S-CVI) - the universal agreement (S-CVI-UA) and its more liberal variant, the average agreement (S-CVI-Ave). S-CVI-UA was defined as the proportion of the items the experts scored as valid (ratings 3 or 4), with the cut-off point: $S-CVI-UA \geq 0.80$ (10). S-CVI-Ave was defined as the average proportion of the items rated 3 or 4, with a cut-off score: $S-CVI-Ave \geq 0.90$ (10). I-CVI was defined as the number of experts providing a rating of 3 or 4/number of experts, with a cut-off score: $I-CVI \geq 0.78$ (9) and automatic item rejection value: $I-CVI < 0.50$ (11). Additionally, we calculated the modified kappa coefficient (k^*) as per Polit (12) to assess chance agreement. The formula for k^* was $(I-CVI - pc) / (1 - pc)$, with pc as the probability of chance occurrence calculated with formula: $[N! / (A!(N-A)!)] \times 0.5^N$, where N is the number of experts and A is the number of experts agreeing on a rating of 3 or 4 (11). The third round of the expert panel involved the same experts from the second round. This time, the questionnaire was administered using the survey app (1KA), and the experts were asked to comment on specific parts and confirm their broad agreement with version MIQ 3.0.

2.2.2 Participants, materials, procedures, data collection, and analysis regarding cognitive testing and pilot study

We conducted cognitive testing with the SOPA MI-based ASBI practitioners as potential respondents to help check the understandability of the items and the questionnaire as a whole. We conducted this testing in two rounds, each proceeding the expert panels’ assessments. In total, we included 10 practitioners, consisting of a family medicine specialist, a specialist in sports medicine, two registered nurses in family medicine practice, two nurses in home care, and four social workers in social work centres. We employed a cognitive interviewing method based on Willis (13), and combined two techniques: think-aloud and verbal probing. Following the reading aloud of the accompanying instructions and questions, respondents were asked to answer questions item by item. They shared in their own words what each question was about, their level of confidence in their understanding, how they interpreted specific terms, their reasoning behind their answers, the difficulty they encountered in responding, and their perception of the comprehensiveness of the response scale. At the end of the interview, we posed additional meta-questions exploring which patients/clients or users they had in mind while responding, whether they provided principle-based answers to any questions, and whether they anticipated answering any questions differently when completing the questionnaire in a conventional

manner. Throughout the process we encouraged the participants to express their thoughts and suggestions, especially when they detected areas for improvement. These interviews were conducted during national COVID restrictions, primarily via telephone or Zoom, and were audio-recorded. The duration of the interviews ranged from 47 to 123 minutes, and due to their length two of them were conducted in two parts. We conducted a preliminary analysis during the interviews, followed by a more in-depth analysis upon reviewing the recordings.

After we aligned the feedback from MI experts and practitioners for version 3.0, we additionally sought comments from respondents in the succeeding pilot study regarding the questionnaire. As with the cognitive testing, the respondents were SOPA MI practitioners. Due to the small sample size ($n=31$) and potential data identifiability, we did not collect further details on the sampled individuals. The pilot version of the questionnaire was administered via a survey app (1KA) in the autumn of 2020. Participation in all steps of the questionnaire validation process was entirely voluntary and without any financial incentives.

3 RESULTS

3.1 Stage one - questionnaire development

We initially identified 19 articles and subsequently excluded four articles either because they focused on self-assessment in patients/clients (14-16) or did not address the practice of MI (17).

Analysis of the remaining articles revealed the following: All of the current self-assessment tools were developed or published after 2003 (18-32), with almost half from 2020 onwards (10, 27-32).

Most of the self-assessment tools focused on evaluating the implementation of MI in a single conducted session (20, 21, 23-27, 29-31). In some cases, these tools were derived from instruments used to assess MI integrity, such as MISC (27) or MITI (20, 31), or from a supervisory tool according to MIA:STEP (21). In some of the other instances, they took the form of checklist-style inventories (27, 28).

Various tools addressed the use of different elements of MI. Almost all of them encompassed selected aspects of the spirit of MI, with many focusing on skills (21-24, 27, 29) and emphasizing strategies for assessing readiness for change (18, 21, 26, 29, 32). The range of relevant items in these tools varied from one (19) to 20 (23). Some items were 'double/triple etc.-barreled' (actually contained two/three etc. different questions within one) (21, 8).

Response scales were often 5-point (18-20, 23, 27, 31) or 3-point (20, 29, 31), but some were 4-point (24, 25), binary (26), 6-point (29), or 7-point (21). The scales measured frequency (18), agreement (23), the extent of behaviour (19-21, 31), or the number of occurrences of behaviour

(20, 31), expertise (24, 25), optimality (27), or capability (29). In three cases, the possible response scales were not described (28, 30, 32).

From our literature review, we generated a pool of 58 items addressing five important aspects of MI: partnership, acceptance, evoking, resisting the righting reflex, and strengthening self-efficacy. Some aspects of MI we did not assess include: focussing, planning, compassion, and developing discrepancy. For practical reasons, we reduced the number of items to 30. We introduced a 7-point frequency scale and included instructions for completing the questionnaire. This marked the creation of version one of the MI questionnaire (MIQ 1.0).

3.2 Stage two - judgment and quantification

The first expert panel round revealed concerns about the clarity of instructions and the understandability of items. Specifically, there were questions about what the period or frequency referred to, whether it was about the total number of times the element was practiced, the number of times in one session, with one or all patients/clients, or the duration when it occurred. Some experts raised concerns about the questionnaire's length and abundant response options, and some questioned the often indifferent neutral middle option. Certain sections were questioned regarding their understandability, and these concerns were given special consideration during the subsequent cognitive testing.

In the first round of cognitive testing all five respondents quickly adapted to the instructions and almost instantly discussed all the required aspects in one flow.

For example:

KT1_1_36-39/1/ (in item P3): "Yes - (reads the question:) How often have you checked if you and the patient (skips the words 'slash client') are working together towards the same goal? (short pause, thinking) How often? Well... this actually refers to, it refers to one patient, if I understand correctly, I would interpret it this way: it refers to one patient over several sessions or encounters (note: it means meetings), and I would understand it as, do I check with the patient at each encounter if we are on the same path (short pause)... yes, I would answer (short pause) 'almost always'. Almost every time the patient came for an encounter, I somehow checked, actually, even between the lines, if we were heading toward the same goal. I would answer 'almost always'."

Moderator: (waits for a moment) "I see, okay, now you've also told me how you came to your thoughts. What do you think of this question - is it difficult/easy, understandable?"

KT1_1_36-39(1): (short pause) "I find this question quite okay. It's fine with me. Good."

Moderator: "And what about the appropriateness of the answers, are they fine? The options, are they okay?"

KT1_1_36-39(1): "Yes, 'never' is out, well, 'almost every time', yes, you kind of refresh or check at almost every encounter if we are both working toward the same goal. I could choose 'frequently', well, either 'frequently' or 'almost always' I would choose."

Moderator: "I see, what would you choose?"

KT1_1_36-39(1): (pause) "Now, if there were only 'never', 'sometimes', 'always', I would choose 'sometimes', well, now, because I have two more sub-questions, 'frequently' and 'almost always'. yes, I chose 'almost always.'"

Some answers were based less on actual experience (or memory of it) and were more principle-based or considered less thoughtfully. This was primarily the case for some items related to the spirit of MI, particularly partnership and acceptance. Items containing the verbs "ask" and "tell" were affected to a lesser extent. At times different respondents or even the same respondent had particular patients/clients in mind. Respondents showed a good general understanding of the optimal practice of MI elements. Some testers liked the multiple response options, some found them unnecessary. Typically, respondents tended to select the middle answer with less

consideration, and they did not encounter difficulty in choosing an adjacent option when prompted.

Using the respondents' answers, we adapted the instructions to be more precise and direct in completing the questionnaire based on actual experience over principle-based answers. We added adverbial or adjectival emphasis to certain words and underlined them (e.g., actively strive). We also removed the middle option for answers. This resulted in the creation of version two of the MI questionnaire (MIQ 2.0).

In the second expert panel round the indices and the modified kappa coefficient indicated that some experts found understandability problematic with regard to the elements of partnership, acceptance and resisting the righting reflex, and relevance in element evoking according to the S-CVI-UA value. However, no item had any of the four categories indices with values lower than 0.50, at which point an item would automatically be removed, as indicated by Halek (11). As suggested in the literature (9, 10), they were instead taken into special consideration for further adaptation and/or testing. Detailed values of the indices and k*s in all four categories are presented in Tables 1, 2 and 3.

Table 1. The content validity of the measurement instrument as a whole and by specific MI elements, with the universal agreement of experts (S-CVI-UA).

MI scale/element	S-CVI-UA ¹			
	RELEVANCE of the question	UNDERSTANDABILITY of the question	COMPLETENESS of response options	MEANINGFULNESS of response options
Scale as a whole	0.93	0.67	0.93	0.85
Partnership	1.00	0.33	1.00	0.67
Acceptance	0.80	0.40	0.80	0.80
Evoking	0.75	1.00	1.00	1.00
Resisting the righting reflex	1.00	0.67	0.83	0.83
Strengthening self-efficacy	1.00	1.00	1.00	1.00

Legend: ¹ S-CVI-UA = the proportion of the items the experts scored as valid (ratings 3 or 4); cut-off point: S-CVI- UA \geq 0.80 (10)

Table 2. The content validity of the measurement instrument as a whole and by specific MI elements, with the universal agreement of experts (S-CVI-UA).

MI scale/element	S-CVI-Ave ¹			
	RELEVANCE of the question	UNDERSTANDABILITY of the question	COMPLETENESS of response options	MEANINGFULNESS of response options
Scale as a whole	0.99	0.93	0.98	0.98
Partnership	1.00	0.86	0.94	0.94
Acceptance	0.97	0.73	0.97	0.97
Evoking	0.96	1.00	1.00	1.00
Resisting the righting reflex	1.00	0.94	0.95	0.97
Strengthening self-efficacy	1.00	0.95	1.00	1.00

Legend: ¹ S-CVI- Ave = the average proportion of the items rated 3 or 4; cut-off score: S-CVI-Ave \geq 0.90 (10)

Table 3. Values of the validity index for individual items (I-CVI) and the modified kappa coefficient (k*) for 27 items.

MI element	Item code* and content	RELEVANCE of the question				UNDERSTANDABILITY of the question				COMPLETENESS of response options				MEANINGFULNESS of response options			
		N ¹ _(exp3-4)	I-CVI ²	p _c ³	k ⁴	N ¹ _(exp3-4)	I-CVI ²	p _c ³	k ⁴	N ¹ _(exp3-4)	I-CVI ²	p _c ³	k ⁴	N ¹ _(exp3-4)	I-CVI ²	p _c ³	k ⁴
PARTNERSHIP	P1 make P/C ⁵ feel comfortable	6	1.00	0.000	1.00	5	0.83	0.094	0.67	5	0.83	0.094	0.67	5	0.83	0.094	0.67
	P2 being supportive	6	1.00	0.000	1.00	5	0.83	0.094	0.67	5	0.83	0.094	0.67	5	0.83	0.094	0.67
	P3 working together	6	1.00	0.000	1.00	6	1.00	0.000	1.00	6	1.00	0.000	1.00	6	1.00	0.000	1.00
	P4 P/C's input	6	1.00	0.000	1.00	4	0.67	0.234	0.33	6	1.00	0.000	1.00	6	1.00	0.000	1.00
	P5 incorporate P/C's ideas	6	1.00	0.000	1.00	6	1.00	0.000	1.00	6	1.00	0.000	1.00	6	1.00	0.000	1.00
	P6 be there in case P/C changes their mind	6	1.00	0.000	1.00	5	0.83	0.094	0.67	6	1.00	0.000	1.00	6	1.00	0.000	1.00
ACCEPTANCE	A1 P/C's view is relevant	6	1.00	0.000	1.00	6	1.00	0.000	1.00	6	1.00	0.000	1.00	6	1.00	0.000	1.00
	A2 strive to understand	6	1.00	0.000	1.00	5	0.83	0.094	0.67	5	0.83	0.094	0.67	5	0.83	0.094	0.67
	A3 P/C's choice to change	6	1.00	0.000	1.00	5	0.83	0.094	0.67	6	1.00	0.000	1.00	6	1.00	0.000	1.00
	A4 respect P/C's decision	6	1.00	0.000	1.00	6	1.00	0.000	1.00	6	1.00	0.000	1.00	6	1.00	0.000	1.00
	A5 P/C's personal growth	5	0.83	0.094	0.67	4	0.67	0.234	0.33	6	1.00	0.000	1.00	6	1.00	0.000	1.00
EVOKING	E2 P/C's own reasons	6	1.00	0.000	1.00	6	1.00	0.000	1.00	6	1.00	0.000	1.00	6	1.00	0.000	1.00
	E3 P/C's own strategies	6	1.00	0.000	1.00	6	1.00	0.000	1.00	6	1.00	0.000	1.00	6	1.00	0.000	1.00
	E4 encourage P/C's thinking	5	0.83	0.094	0.67	6	1.00	0.000	1.00	6	1.00	0.000	1.00	6	1.00	0.000	1.00
	E5 P/C's inner strenghts and sources	6	1.00	0.000	1.00	6	1.00	0.000	1.00	6	1.00	0.000	1.00	6	1.00	0.000	1.00
RESISTING THE RIGHTING REFLEX	R1* explaining without first exploring	6	1.00	0.000	1.00	6	1.00	0.000	1.00	6	1.00	0.000	1.00	6	1.00	0.000	1.00
	R2* talking about own knowing	6	1.00	0.000	1.00	5	0.83	0.094	0.67	5	0.83	0.094	0.67	5	0.83	0.094	0.67
	R3* reasons without permission and inquire	6	1.00	0.000	1.00	6	1.00	0.000	1.00	6	1.00	0.000	1.00	6	1.00	0.000	1.00
	R4* ideas without permission and inquire	6	1.00	0.000	1.00	6	1.00	0.000	1.00	6	1.00	0.000	1.00	6	1.00	0.000	1.00
	R5* talk P/C into	6	1.00	0.000	1.00	5	0.83	0.094	0.67	6	1.00	0.000	1.00	6	1.00	0.000	1.00
	R6 suggestions after permission and inquire	6	1.00	0.000	1.00	6	1.00	0.000	1.00	6	1.00	0.000	1.00	6	1.00	0.000	1.00
STRENGTHENING SELF-EFFICACY	S1 ask about confidence	6	1.00	0.000	1.00	6	1.00	0.000	1.00	6	1.00	0.000	1.00	6	1.00	0.000	1.00
	S2 ask about needed	6	1.00	0.000	1.00	6	1.00	0.000	1.00	6	1.00	0.000	1.00	6	1.00	0.000	1.00
	S3 P/C's past experiences	6	1.00	0.000	1.00	6	1.00	0.000	1.00	6	1.00	0.000	1.00	6	1.00	0.000	1.00
	S5 affirmations	6	1.00	0.000	1.00	6	1.00	0.000	1.00	6	1.00	0.000	1.00	6	1.00	0.000	1.00
	S6 change talk	6	1.00	0.000	1.00	5	0.83	0.094	0.67	6	1.00	0.000	1.00	6	1.00	0.000	1.00
	S8 other resources	6	1.00	0.000	1.00	6	1.00	0.000	1.00	6	1.00	0.000	1.00	6	1.00	0.000	1.00

Legend:

* reverse scaling

¹ N(exp3-4) = number of experts providing a rating of 3 or 4

² I-CVI (content validity index) = number of experts providing a rating of 3 or 4/number of experts; cut-off score: I-CVI≥0.78 (9); automatic item rejection: I-CVI<0.50 (11)

³ pc (probability of chance occurrence) = [N!/A!(N-A)!] x 0,5N N = number of experts; A = number of experts agreeing on a rating of 3 or 4 (11)

⁴ k* (modified kappa) = (I-CVI-pc)(1-pc)

⁵ P/C = patient/client

In the second round of cognitive testing, the respondents properly understood both items that were considered potentially problematic in terms of understandability by the expert panel. For example, the expression “personal growth”, considered too broad and not understandable by two panel experts in round two, was consistently viewed by respondents in the second round of cognitive testing as one’s general ability to change one’s way of thinking and behaving, to undergo the necessary behaviour change, or to stop drinking (excessively). In this round the respondents also demonstrated appropriate knowledge regarding the optimal practice of different MI elements and remembered different patients/clients and situations. However, they provided fewer principle-based answers (although some instances still occurred, again in the partnership and acceptance subscales) and relied more on their memory of actual situations.

Based on insights and suggestions from the second round of the expert panel and the cognitive testing we made changes to some expressions, and divided some items into two separate questions, made further improvements to the instructions, and created the third version of the MI questionnaire (MIQ 3.0).

This final version of the questionnaire was then approved by the expert panel in the third round, and no further comments were received from respondents during the questionnaire piloting.

4 DISCUSSION

The main aim of this study was to develop a comprehensive self-assessment questionnaire about practicing MI in conducting ASBI and to test its content validity. We used an iterative process involving a literature review, expert panel method and cognitive testing. This resulted in a content-valid 30-item long self-assessment questionnaire with a 6-point response scale exploring five elements of MI practice when conducting ASBI.

Based on our review, previous studies have neither generated nor used a comprehensive and content-valid self-assessment questionnaire for MI practitioners that can be used for assessing MI practice over extended time periods - e.g. weeks or months. One study, however (18), did ask practitioners two MI questions (out of 39) about past practices in smoking cessation counselling. These two items focused on the self-assessment of the importance of change and confidence in making the change. This earlier questionnaire showed good content validity and internal consistency (18), and we included these aspects of those items in our questionnaire.

The expert panel’s opinion can be analysed in different ways (e.g. 34, 35). In our case this involved calculating different content indexes in four content categories,

allowing us to analyse the experts’ opinions very systematically and to pinpoint exactly where the potential problem was and what we needed to do about it. The otherwise acceptable to high or even optimal values of the CVIs (I-CVI, S-CVI-UA/Ave) and k^* were most negatively affected by two items due to the expert panel’s concerns about understandability. Due to the fact that none of the items had their index value lower than 0.5, they were not automatically rejected. Similar to Halek et al. (11) and Carli et al. (34) in such cases, these items were further tested with potential respondents.

In the iterative process of cognitive interviewing, the respondents understood both previously problematised items by the expert correctly and so the questions remained. In some other questions, at first some expressions were less understandable, and some items were answered in a more principle-based manner. These items were adjusted and in the subsequent testing the questions were understood accurately and were answered more based on the memory of the respondents’ actual experiences. Similarly, Robinson et al. (36) succeeded in enhancing the understandability of the questionnaire substantially by conducting this iterative process. In this way, our results confirmed the value of cognitive interviewing as a powerful tool for gaining insight into the thought process of the respondents and for improving the understandability of the questionnaire (as per Willis) (13).

Our study has potential limitations that need to be addressed. Firstly, we focused on five MI elements, a mixture of selected aspects of the MI spirit, processes and principles, whilst leaving some of the aspects of these as well as skills, strategies and techniques out. This is not unique to our questionnaire, but is rather a common feature of other questionnaires and MI assessment tools which also cover different selected aspects of the MI spirit and/or different selected behaviours as stated earlier in this article. Which MI practice variables are selected and how they are captured varies at least to a certain degree. As per Moyers et al. (6), it is acceptable not to include some aspects to reduce the complexity of the tool whilst also being clear about those elements or aspects that are included.

Next, according to our cognitive testing results, respondents might answer some questions in a more principle-based manner and/or less thoughtfully, usually more with those items seeking to capture the spirit of MI. We tried to reduce this tendency by adding instructions about the importance of answering according to actual personal experience rather than the professional ideal, and emphasising the practical value of completing the questionnaire in a manner which encourages reflection on one’s personal MI practice. Whilst principle-based answering may reflect a respondent’s difficulty in assessing their personal performance, Beckman et al.

(31) comment on the effect of metacognition, as (self-) estimates may become more accurate during repeated testing and subjective ratings become more aligned with the objective ones. Regular use of check-lists and supervision may further add to this alignment of subjective and objective performance rating (31), especially when divergence encourages reflection and deliberate practice. Nonetheless, combining self-assessment instruments with objective ratings of performance may be optimal (31).

The questionnaire we developed is not short. The MI questionnaires mentioned earlier have up to 20 items, while ours has 30, and some of the experts felt that the length might lower the response rates. However, the respondents in the cognitive testing part of this study did not comment on the questionnaire being too long. As per Robinson (33), to fully capture the richness of multidimensional variables, a larger number of items is required. In our case, the five MI elements we chose to incorporate could potentially mean five different dimensions of the questionnaire. Having approximately six items per element before testing the psychometric properties and potentially needing to narrow down the number of items per element/dimension/subscale to three, as the generally recommended minimum (33), makes this a rational decision.

Finally, some of the items are alcohol-risk-factor-specific, and the language of the questionnaire is Slovenian. These specifics call for additional content validity testing when planning to use the instrument in a broader context and/or different languages.

There have been very few published studies on self-assessment of practicing MI. This is a rather young research field, as the majority of identified studies were published after 2015, half of them after 2020. Our study focused on the content validity of the questionnaire, leaving it open for further validation processes, including testing its psychometric properties, as in, for example, Sočan et al. (37).

5 CONCLUSIONS

To the best of our knowledge, this is the only study in the MI research field that has deployed such a rigorous and comprehensive procedure for establishing the content validity of a self-assessment questionnaire. The questionnaire's final version demonstrates appropriate content validity and is ready for testing its psychometric properties. With regard to reducing its length, we suggest the first items to be removed are those with a potentially higher likelihood of principle-based responses.

ACKNOWLEDGMENTS

We would like to thank all ten MI experts for their opinions on the content validity of the questionnaire as well as all the ten MI-based ASBI practitioners in the TRATAC (Slovenian: SOPA) project who assessed the questionnaire's understandability in the cognitive testing, and 31 who collaborated it in the pilot study.

CONFLICTS OF INTEREST

The authors declare that there are no conflicts of interest.

FUNDING

This research was co-financed by the European Commission (European Social Fund) and Republic of Slovenia within the Operational Programme for the Implementation of the EU Cohesion Policy 2014-2020.

ETHICAL APPROVAL

The study was approved by the National Medical Ethics Committee of the Republic of Slovenia (No. 0120-246/2018/21). All participants gave informed consent prior to study participation.

AVAILABILITY OF DATA AND MATERIALS

The data and materials are securely stored at the facilities of Slovenia's National Institute of Public Health. Further information on the questionnaire's development stages can be made available by the first author upon request.

ORCID

Tadeja Hočvar:
<https://orcid.org/0009-0000-2713-9801>

Tim Anstiss:
<https://orcid.org/0000-0002-7762-532X>

Danica Rotar Pavlič:
<https://orcid.org/0000-0001-7575-3195>

REFERENCES

1. Miller WR, Rollnick S. *Motivational interviewing: Helping people change* 3rd ed. New York: Guilford Press; 2013.
2. DiClemente CC, Corno CM, Graydon MM, Wiprovnick AE, Knoblach DJ. Motivational interviewing, enhancement, and brief interventions over the last decade: A review of reviews of efficacy and effectiveness. *Psychol Addict Behav*. 2017; 31(8):862-887. doi: 10.1037/adb0000318.
3. Miller WR. *Motivational interviewing skill code*. Unpublished document; 2000.

4. Moyers T, Martin T, Catley D, Harris KJ, Ahluwalia JS. Assessing the integrity of motivational interviewing interventions: Reliability of the motivational interviewing skills code. *Behav Cogn Psychother*. 2003;31(2):177-184. doi: 10.1017/S1352465803002054.
5. Amrhein P, Miller WR, Moyers T, Ernst D. Manual for the motivational interviewing skill code (MISC). Department of Psychology Faculty Scholarship and Creative Works; 27 [Internet]. 2008 [cited 2023 Aug 8]. Available from: <https://digitalcommons.montclair.edu/psychology-facpubs/27>
6. Moyers TB, Martin T, Manuel JK, Miller, WR, Ernst D. Revised global scales: Motivational Interviewing Treatment Integrity 3.1.1 (MITI 3.1.1). New Mexico: Center on Alcoholism, Substance Abuse, and Addictions, The University of New Mexico [Internet]. 2010 [cited 2023 Aug 8]. Available from: <https://www.motivationalinterviewing.org/sites/default/files/MITI%203.1.pdf>
7. Moyers TB, Rowell, LN, Manuel JK. The Motivational Interviewing Treatment Integrity code (MITI 4): Rationale, preliminary reliability and validity. *J Subst Abuse Treat*. 2016;65:36-42. doi: 10.1016/j.jsat.2016.01.001.
8. Northwest Frontier Addiction Technology Transfer Center, Oregon Health and Science University. Motivational interviewing assessment: Supervisory tools for enhancing proficiency (MIA: STEP). Salem: Northwest Frontier Addiction Technology Transfer Center, Oregon Health and Science University. [Internet]. 2006 [cited 2023 Aug 8]. Available from: <https://motivationalinterviewing.org/sites/default/files/mia-step.pdf>
9. Lynn MR. Determination and quantification of content validity. *Nurs Res*. 1986; 35:382-385. doi: 10.1097/00006199-198611000-00017.
10. Polit DF, Beck CT. The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Res Nurs Health*. 2006;29:489-497. doi: 10.1002/nur.20147.
11. Halek M, Holle D, Bartholomeyczik S. Development and evaluation of the content validity, practicability and feasibility of the Innovative dementia-oriented Assessment system for challenging behaviour in residents with dementia. *BMC Health Serv Res*. 2017;17(1):554. doi: 10.1186/s12913-017-2469-8.
12. Polit DF, Beck CT, Owen SV. Is the CVI an acceptable indicator of content validity? Appraisal and recommendations *Res Nurs Health*. 2007;30:459-467. doi: 10.1002/nur.20199.
13. Willis GB. *Cognitive interviewing*. Thousand Oaks: SAGE Publications; 2005.
14. Krampe H, Salz AL, Kerper LF, Krannich A, Schnell T, Wernecke KD, et al. Readiness to change and therapy outcomes of an innovative psychotherapy program for surgical patients: Results from a randomized controlled trial. *BMC Psychiatry*. 2017;17(1):417. doi: 10.1186/s12888-017-1579-5.
15. Freyer-Adam J, Baumann S, Bischof G, Staudt A, Goeze C, Gaertner B, et al. Social equity in the efficacy of computer-based and in-person brief alcohol interventions among general hospital patients with at-risk alcohol use: A randomized controlled trial. *JMIR Ment Health*. 2022;9(1):e31712. doi: 10.2196/31712.
16. Bredie SJ, Fouwels AJ, Wollersheim H, Schippers GM. Effectiveness of nurse based motivational interviewing for smoking cessation in high risk cardiovascular outpatients: A randomized trial. *Eur J Cardiovasc Nurs*. 2011;10(3):174-179. doi: 10.1016/j.ejcnurse.2010.06.003.
17. Ramseier CA, Hübschi C, Crnić T, Woelber JP. Implementation of a communication curriculum in undergraduate dental education-students' opinions during a 5-year development phase. *Eur J Dent Educ*. 2023 Aug 9. doi: 10.1111/eje.12940.
18. Scal P, Hennrikus D, Ehrlich L, Ireland M, Borowsky I. Preparing residents to counsel about smoking. *Clin Pediatr (Phila)*. 2004;43(8):703-708. doi: 10.1177/000992280404300803.
19. Coyle K, Carcone AI, Butame S, Pooler-Burgess M, Chang J, Naar S. Adapting the self-assessment of contextual fit scale for implementation of evidence-based practices in adolescent HIV settings. *Implement Sci Commun*. 2022;3(1):115. doi: 10.1186/s43058-022-00349-4.
20. Mullin DJ, Saver B, Savageau JA, Forsberg L, Forsberg L. Evaluation of online and in-person motivational interviewing training for healthcare providers. *Fam Syst Health*. 2016; 34(4):357-366. doi: 10.1037/fsh0000214.
21. Widder R. Learning to use motivational interviewing effectively: Modules. *J Contin Educ Nurs*. 2017;48(7):312-319. doi: 10.3928/00220124-20170616-08.
22. Stenov V, Wind G, Skinner T, Reventlow S, Hempler NF. The potential of a self-assessment tool to identify healthcare professionals' strengths and areas in need of professional development to aid effective facilitation of group-based, person-centered diabetes education. *BMC Med Educ*. 2017;17(1):166. doi: 10.1186/s12909-017-1003-3.
23. Vallabhan MK, Kong AS, Jimenez EY, Summers LC, DeBlick CJ, Feldstein Ewing SW. Training primary care providers in the use of motivational interviewing for youth behavior change. *Res Theory Nurs Pract*. 2017;31(3):219-232. doi: 10.1891/1541-6577.31.3.219.
24. Muzyk AJ, Tew C, Thomas-Fannin A, Dayal S, Maeda R, Schramm-Sapyta N, et al. Utilizing Bloom's taxonomy to design a substance use disorders course for health professions students. *Subst Abuse*. 2018; 39(3):348-353. doi: 10.1080/08897077.2018.1436634.
25. Galal S, Vyas D, Mayberry J, Rogan EL, Patel S, Ruda S. Use of standardized patient simulations to assess impact of motivational interviewing training on social-emotional development. *Pharmacy (Basel)*. 2018;6(3):65. doi: 10.3390/pharmacy6030065.
26. Boykan R, Blair R, Baldelli P, Owens S. Using motivational interviewing to address tobacco cessation: Two standardized patient cases for pediatric residents. *MedEdPORTAL*. 2019;15:10807. doi: 10.15766/mep_2374-8265.10807.
27. Spangler JG, Shull CN, Hildebrandt CA, Jones KB, Brewer AL, Knudson MP, et al. Opioid use disorder and assessment of patient interactions among family medicine residents, medical students, and physician assistant students. *MedEdPORTAL*. 2020;16:11012. doi: 10.15766/mep_2374-8265.11012.
28. Sadasivam RS, Kamberi A, DeLaughter K, Phillips B, Williams JH, Cutrona SL, et al. Secure asynchronous communication between smokers and tobacco treatment specialists: Secondary analysis of a web-assisted tobacco intervention in the QUIT-PRIMO and national dental PBRN networks. *QUITPRIMO; National Dental PBRN Collaborative Group. J Med Internet Res*. 2020;22(5):e13289. doi: 10.2196/13289.
29. Arnett MC, Evans MD, Stull C. Dental hygiene students' perceptions regarding the importance of and confidence with using brief motivational interviewing during HPV patient counseling. *J Dent Hyg*. 2022;96(2):50-58.
30. Arnett MC, Evans MD, Stull CL. Brief motivational interviewing: Evaluation of a skills-based education program. *J Dent Hyg*. 2022;96(4):46-56.
31. Beckman M, Lindqvist H, Öhman L, Forsberg L, Lundgren T, Ghaderi A. Correspondence between practitioners' self-assessment and independent motivational interviewing treatment integrity ratings. *Front Psychol*. 2022;13:890579. doi: 10.3389/fpsyg.2022.890579.
32. Fezza GC, Sansone S, Nolan RP. Therapeutic components of digital counseling for chronic heart failure. *Front Psychiatry*. 2022;13:888524. doi: 10.3389/fpsyg.2022.888524.
33. Robinson MA. Using multi-item psychometric scales for research and practice in human resource management. *Hum Resour Manage*. 2018;57:739-750. doi: 10.1002/hrm.21852.
34. Carli T, Košnik M, Zaletel-Kragelj L, Burazeri G, Kukec A. The APISS questionnaire: A new tool to assess the epidemiology of systemic allergic reactions to bee venom in beekeepers. *Zdr Varst*. 2023;62(3):137-144. doi: 10.2478/sjph-2023-0019.
35. Slosar L, Puš K, Marušič U. Validation of the Slovenian version of the Movement Imagery Questionnaire for Children (MIQ-C): A measurement tool to assess the imagery ability of motortasks in children. *Zdr Varst*. 2023;62(3):113-120. doi: 10.2478/Sjph-2023-0016.
36. Robinson KM, Scherer AM, Nishimura TE, Laroche HH. Value of cognitive interviewing in the development of the weight stigma in healthcare inventory. *Patient Educ Couns*. 2023;113:107767. doi: 10.1016/j.pec.2023.107767.
37. Sočan G, De Boer D, Murko E, Kralj M, Ropret N, Zaletel M. Psychometric validation of an instrument for measuring patient experiences with outpatient healthcare. *Zdr Varst*. 2023;62(3):153-161. doi: 10.2478/sjph-2023-0021.