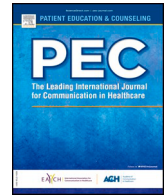




Contents lists available at ScienceDirect

Patient Education and Counseling

journal homepage: www.elsevier.com/locate/pec

Review article

What do we know about written assessment of health professionals' communication skills? A scoping review

Noelle Junod Perron^{a,*}, Peter Pype^{b,1}, Marc van Nuland^c,
 Maria Magdalena Bujnowska-Fedak^d, Marcela Dohms^e, Geurt Essers^f, Ragnar Joakimsen^g,
 Zoi Tsimtsiou^h, Claudia Kiesslingⁱ

^a Unit of Development and Research in Medical Education, Geneva Faculty of Medicine and Institute of Primary Care, Geneva University Hospitals, Geneva, Switzerland

^b Department of Public Health and Primary Care, Ghent University, Ghent, Belgium

^c Academic Center for General Practice, Leuven University, Leuven, Belgium

^d Department of Family Medicine, Wrocław Medical University, Wrocław, Poland

^e Radboud University Medical Center, The Netherlands

^f Network of GP Training Programs in the Netherlands, Utrecht, The Netherlands

^g Department of Clinical Medicine, Faculty of Health Sciences, UiT The Arctic University of Norway and Department of Internal Medicine, University Hospital of North Norway, Tromsø, Norway

^h Department of Hygiene, Social-Preventive Medicine and Medical Statistics, School of Medicine, Aristotle University of Thessaloniki, Thessaloniki, Greece

ⁱ Personal and Interpersonal Development in Health Care Education, Witten/Herdecke University, Witten, Germany

ARTICLE INFO

Article history:

Received 2 March 2021

Received in revised form 26 August 2021

Accepted 6 September 2021

Available online xxxx

Keywords:

Communication skills

Health professionals

Knowledge

Written assessment

ABSTRACT

Objective: The aim of this scoping review was to investigate the published literature on written assessment of communication skills in health professionals' education.

Methods: Pubmed, Embase, Cinahl and Psychinfo were screened for the period 1/1995–7/2020. Selection was conducted by four pairs of reviewers. Four reviewers extracted and analyzed the data regarding study, instrument, item, and psychometric characteristics.

Results: From 20,456 assessed abstracts, 74 articles were included which described 70 different instruments. Two thirds of the studies used written assessment to measure training effects, the others focused on the development/validation of the instrument. Instruments were usually developed by the authors, often with little mention of the test development criteria. The type of knowledge assessed was rarely specified. Most instruments included clinical vignettes. Instrument properties and psychometric characteristics were seldom reported.

Conclusion: There are a number of written assessments available in the literature. However, the reporting of the development and psychometric properties of these instruments is often incomplete. Practice implications written assessment of communication skills is widely used in health professions education. Improvement in the reporting of instrument development, items and psychometrics may help communication skills teachers better identify when, how and for whom written assessment of communication should be used.

© 2021 The Authors. Published by Elsevier B.V.
 CC-BY-NC-ND 4.0

Contents

1. Introduction	3
2. Method	3

* Correspondence to: Unit of Development and Research in Medical Education, Faculty of Medicine, Geneva University and Institute of Primary Care, Geneva University Hospitals, Switzerland.

E-mail address: noelle.junod@hcuge.ch (N.J. Perron).

¹ P. Pype and M. Van Nuland contributed equally

<https://doi.org/10.1016/j.pec.2021.09.011>

0738-3991/© 2021 The Authors. Published by Elsevier B.V.
 CC-BY-NC-ND 4.0

Please cite this article as: N.J. Perron, P. Pype, M. van Nuland et al., What do we know about written assessment of health professionals' communication skills? A scoping review, Patient Education and Counseling, <https://doi.org/10.1016/j.pec.2021.09.011>

2.1.	Identifying the research question	3
2.2.	Identifying relevant studies	3
2.3.	Selecting the studies	4
2.4.	Charting the data	4
2.5.	Collating, summarizing and reporting the results	4
3.	Results	4
3.1.	Study characteristics	4
3.1.1.	Study aims	4
3.1.2.	Study designs	4
3.1.3.	Study populations	4
3.2.	Instrument characteristics	5
3.2.1.	Composition and purpose of instruments	6
3.2.2.	Length of instruments /number of items in a table	6
3.2.3.	Instrument development	6
3.2.4.	Communication topic being assessed	6
3.2.5.	Type of knowledge assessed	7
3.2.6.	Scoring activities	7
3.3.	Item characteristics	7
3.3.1.	Stimulus	7
3.3.2.	Response format	7
3.4.	Psychometric properties	7
3.4.1.	Instrument or item properties	8
3.4.2.	Reliability	8
3.4.3.	Validity	8
3.5.	Instrument availability	8
4.	Discussion and conclusion	8
4.1.	Discussion	8
4.2.	Conclusion	9
4.3.	Practical implications	9
	CRedit authorship contribution statement	9
	Funding	10
	Declaration of Competing Interest	10
	Acknowledgement	10
	Appendix A : Search strategy for Pubmed	11
	Appendix B : Differences between behavioural and knowledge-oriented instructions	11
	Appendix C	11
	Appendix D Examples of stimuli and responses (modified examples extracted from the articles)	11
	References	12

1. Introduction

Effective communication skills are recognized to be core competences for healthcare professionals. Although there is no clear consensus across the professions regarding what constitutes effective skills for communication [1], it has been shown that communication skills can be successfully trained among different groups of health professionals [2–5] and in specific contexts [6–10]. Several countries have integrated communication skills into their requirements for healthcare education or in their safety and quality of care standards [11–18]. Documents and articles provide evidence and guidance for the implementation of communication skills training, for specific professions, at different milestones and for specific topics [19–23].

Numerous assessment tools and methods can be used to assess communication skills within an educational program. They are usually categorized according to the level of competence they aim to assess and the purpose and desired outcomes of the teaching program [24–26]. The “shows” and “does” levels of clinical competence are - according to the Miller pyramid [27] - best assessed with OSCEs [28–31] or in the workplace [26,32,33] because they allow for the observation of communication behaviors. The “knows” and “knows how” levels most commonly are assessed through factual tests or context-based tests such as MCQ, essays, or oral examination [32,34,35].

The assessment of the knowledge/cognitive component of communicative competence varies. A recent literature review listed four learning outcomes for communication skills [36]: knowledge – one’s

understanding of information through which incoming data and/or experiences are processed and recorded; content – what is communicated; process – how one communicates; and perceptual skills – awareness of self and others and how that impacts communication. The authors concluded that knowledge is a component of the first phases of scaffolding for learning and assessment and represents 20% of all learning outcomes for communication skills [36].

According to Hargie, an important first step for learning a social skill is the acquisition of conceptual, conditional and procedural knowledge, which is then encapsulated to the application of skills [37]. Communication skills are thought to be controlled by psychological mechanisms incorporating cognitive schemata or scripts for action that help construct interpretations of situations [38]. Effective communicative actions, which result from these interpretations therefore are not “right” or “wrong” actions but range from an evidence-based or expert-based corridor of good practice which allows room for personal scope and contextual adaptation. The actions encompass several steps such as understanding what is happening in the situation, considering the range of options open to the participants, and analyzing the consequences of these options [39]. From this perspective, it seems to be helpful for novice learners to identify examples of effective and ineffective communication behaviors, to study examples, to recognize the consequences for the patient, to demonstrate situational awareness and provide reasonable alternative responses [37,38].

Several models have been developed to classify the assessment of cognitive processes or knowledge dimensions. The most well-known model is Bloom’s taxonomy which provides definitions of six major

categories of cognition: knowledge, comprehension, application, analysis, synthesis and evaluation [40]. Krathwohl divided the cognitive process dimensions into factual, conceptual, procedural and metacognitive knowledge [41].

In order to assess knowledge, written assessment requires a stimulus and a response format. Stimuli are classified to be either context-rich or context-poor according to the type of knowledge being measured [35]. A context-rich stimulus may include patient vignettes, scenarios or videos of clinical encounters. It requires learners to apply knowledge to concrete clinical situations and leads to problem-solving. Some are behavior-oriented (e.g. what would you do next) – others are more knowledge-oriented (e.g. which communication techniques provide structure to the patient at this point of the conversation?) [42]. Context –rich stimuli test complex cognitive processes that are characteristic of clinical practice and ask the learner to weigh different units of information against each other. Context-poor stimuli do not include a contextualized vignette or scenario and focus on questions about recall of facts or procedural/conceptual knowledge. It often includes verbs such as “name”, “identify”, “define”, or “explain”.

Response formats in written assessments usually use selected or constructed responses [35]. Selected responses provide a set of predefined answer alternatives. They are typically responses to multiple choice questions and endorse several different formats: true-false, best answer, more than one correct answer etc. Constructed responses invite test takers to write down short answers such as phrases or short sentences, or long answers and essays. Research has shown that it is the content of the lead-in question or stimulus more than the response format that determines what the question tests [43]. The criteria for choosing an instrument and the question types requires consideration of the teaching and learning goals, the reliability and validity of the measure, the educational impact and resource implications, and finally the feasibility of administration of the assessment [35]. Reliability refers to the accuracy of a score on a test [44]. Validity refers to whether the instrument actually tests what it is supposed to test [43,45].

In this scoping review, we aimed to investigate the extent, range and nature of published research activity regarding the use of written assessments of the cognitive component of communication skills in health professionals' education [46]. We were interested in gaining more insight into the aims and designs of studies which included written assessment of communication skills; the communication topics addressed and the learners' profiles. We aimed to evaluate whether the authors described how the instrument was developed, what type of stimulus and response formats were used to test the different cognitive dimensions of communication and whether the psychometric characteristics were reported.

2. Method

We followed the five-step approach described by Arksey and O'Malley to conduct a rigorous scoping review [46]: 1) identifying the research question 2) identifying relevant studies 3) study selection 4) charting the data 5) collating, summarizing and reporting the results. The process was dynamic and iterative. Each step was discussed with the group of investigators and documented in minutes as part of the quality assurance of our work.

2.1. Identifying the research question

Our scoping review focused on answering the question: What do we know about written assessment of communication skills in health professionals' education?

2.2. Identifying relevant studies

The primary search strategy (Pubmed) was developed with all authors and reviewed by an academic librarian. The search strategy used four categories of key terms such as 1) communication, 2) educational measurement/written assessment, 3) computer simulation/computer/video recording/video-based test measurement/paper-based, 4) knowledge/cognition/clinical competence/skill. We tested the relevance and accuracy of the selected key terms by ensuring that key studies published on this topic between 2000 and 2016 were included in the search results. We adapted the search string to the thesaurus of three other databases, namely Embase, Cinahl and Psycinfo with the help of academic librarians (Appendix A). We used EndNote to collect and organize the references. Inclusion criteria were studies published from 1.1.1995 until 2.7.2020, written in English or any other language spoken by the research team (Dutch, French, German, Greek, Norwegian, Polish, and Portuguese). The studies of interest focused on learners in the health professions at the undergraduate, postgraduate or continuous training levels, reporting an educational assessment (test, exam, measurement) and using or aiming at using empirical data to measure the cognitive abilities of individual learners in the field of clinical communication (patient-provider communication). Study designs included protocols, the development, usage and/or validation of a written assessment, interventions, correlation studies, pre and/or post interventions or those which compared different types of interventions. Formats were any stimulus item followed by written answers. Outcome measures included the assessment of learners' written answers indicating cognitive abilities according to Miller's pyramid, Bloom's or Krathwohl's taxonomy [27,40,41]. Using a snowball approach, we added articles of the reference lists if they met the inclusion criteria mentioned above and were not listed in the initial search.

Exclusion criteria were studies focusing on patients, pupils, simulated patients or teachers, using designs such as needs assessments, surveys on patient satisfaction, reporting exclusively outcome measures such as the performance of skills, behaviors, attitudes, self-reported perceptions. Studies assessing knowledge regarding intra- or interprofessional team communication were excluded since communication with patients is only a small part of interprofessional communication. We did not include grey literature.

We conducted the initial search in March 2019, and a follow-up search in July 2020. The comprehensive search strategies used in each data base are available upon request from the first author.

2.3. Selecting the studies

We imported all titles of our search into Rayyan software and deleted duplicates [47]. A calibration exercise with a random sample of 25 abstracts, conducted by all reviewers helped refine the inclusion and exclusion criteria. Reviewers were divided in blinded working pairs and read the titles and abstracts of the studies obtained from the search strategy to determine their eligibility. Studies were categorized as 'include', 'exclude' and 'unsure'. The results of the selection were then unblinded and the same working pairs discussed screening discrepancies. In case of disagreement, the study was categorized as 'include'. Reviewers then retrieved the full texts of all included and unsure studies and decided whether they should be included or not. The selected studies were distributed among the same working pairs for full review. They applied the strict inclusion/exclusion criteria to determine eligibility.

2.4. Charting the data

Data from all included studies were extracted using a data extraction sheet that included the following fields: study publication

information (author, year of publication, country), study characteristics (aims of the study, study design, participants' number and characteristics (e.g. profession and level of learning), and test description (purpose of the test, availability, topic covered, outcome measured, test development and delivery, number of items, test format - item stimulus, prompt and response formats, scoring, psychometrics and feasibility). The same paired reviewers involved in the study selection, extracted data from the selected studies using an excel sheet and discussed discrepancies if necessary. A review of the extracted data led us to realize that reviewers did not have a sufficient common understanding of both the data definition and wording. A sub-group of four reviewers (CK, MvN, NJP, PP) refined the working definitions of the data we wanted to extract, developed selected responses for each criterion, leaving constructed responses when necessary. They conducted a calibration exercise on eight studies, which improved the data extraction sheet. Once the extraction sheet was finalized, the four reviewers used an online survey tool (Qualtrics) to extract the data of the selected studies [48]. Extraction difficulties or discrepancies were resolved by discussion.

2.5. Collating, summarizing and reporting the results

We obtained mainly quantitative results from the data extraction sheet. To gather further information about the level of assessed knowledge, we categorized the lead-in questions, when described, into either knowledge-oriented instructions ("remember", "understand", "analyze" "evaluate" of the Bloom's taxonomy) or behavior-oriented instructions ("apply" and "create/synthesize"). This dichotomy is commensurate with McDaniel's meta-analysis of situational judgement tests (SJT), which stated that SJT response instructions can be categorized into knowledge and behavioral tendency [31] (Appendix B).

3. Results

Our initial search uncovered 23,012 titles of which 2556 were duplicates. After applying our screening tool, we assessed 169 full text studies and identified 74 studies which met the inclusion criteria and underwent full data extraction (see Fig. 1).

3.1. Study characteristics

The studies described projects from 22 countries: most studies coming from the USA (n = 31), the Netherlands (n = 8), Germany (n = 6), and UK (n = 5). One article reported a study conducted in two countries [49]. The number of included publications increased in the last 25 years with the majority of studies published within the last ten years.

3.1.1. Study aims

A majority of studies focused on the development and evaluation of a training intervention (further on called "intervention studies IS") and used written assessment as a measure of the effect of the intervention (Appendix C). The remaining studies focused on the description, development and/or validation of an assessment approach or instrument (further on called "test development studies TDS"). Two studies included both aims but were classified as test development studies to facilitate the analysis [50,51]. Details are presented in Table 1.

3.1.2. Study designs

The vast majority of studies used a quantitative approach. Most intervention studies (IS) used two measurement points while test development studies (TDS) used mainly one measurement point designs (Table 1). Sixteen studies described a study design which included a control group or the comparison of different intervention groups.

3.1.3. Study populations

The number of participants included into the studies showed a large range from 14 [52] to more than 1000 [53–57]. The mean number of participants was higher in the TDS than in the IS (Table 1). Most studies included participants from one professional background, and undergraduate medical students were the most common group of participants (Table 1). Studies with more than one profession focused mainly on participants in continuous training [58–62].

Three studies provided a description of the instrument without its application to any test-takers (e.g. study protocol, description of instruments without data collection) [63–65].

3.2. Instrument characteristics

For clarity purposes, instrument characteristics will be described separately for test development studies (n = 43) and intervention studies (n = 27). The difference between the number of studies (n = 74) and the number of instruments (n = 70) is explained by the fact that six author groups described the same instrument in two different studies [39,50,51,55,56,66–72] (in blue in Appendix C); two studies described two different instruments in the same article [73,74] (in green in Appendix C).

3.2.1. Composition and purpose of instruments

Written assessment of knowledge of communication skills was usually assessed together with other measures. Details regarding measure combinations are displayed in Table 1. In the TDS, a stand-alone instrument was more often used (50%) than in IS (10%) [50,52,53,75,76] (Appendix C).

The information provided in the studies was often not sufficient to evaluate whether the instrument was used for summative or formative assessment purposes.

3.2.2. Length of instruments /number of items in a table

The number of questions used in the instrument was sometimes difficult to assess since some authors indicated the number of questions, others mentioned the number of vignettes without specifying the number of questions deriving from the vignette and some did not provide any information regarding the items. As far as we could detect, the longest stand-alone instrument had 109 questions [72]. The shortest stand-alone test included three videos [75] or six questions [77].

3.2.3. Instrument development

Two-third of the instruments were developed by the authors (Table 2) based on different development criteria: the use of a theoretical framework/guideline/textbook, a literature search, involvement of a group of developers/teachers, a review process or other strategies (e.g. post hoc analysis). The authors commonly reported the use of only one or two criteria to develop their instrument (Table 2). When one criterion was reported, involvement of a group of developers was the most frequently mentioned (n = 7) [54,60,61,78–81], followed by a literature review (n = 3) [82–85]. For the instruments reporting two criteria [50,55,59,62,68,86–91], involvement of group of developers (n = 10), review process (n = 5) and use of theory (n = 4) were most frequently reported and the most common combination was involvement of a group of developers and review process [29,56,87–89]. For nine instruments, three criteria were reported [39,64,73,75,92–95] and for one instrument, four criteria were reported [96].

The most frequent theoretical frameworks explicitly reported were Miller and Rollnick's model of motivational interviewing [59,77,97–103], COMFORT [62,91,104], Calgary Cambridge Guide [49,105], Kalamazoo Consensus Statement (general communication skills) [75,90], SPIKES model (breaking bad news) [106,107], ACGME (core competencies for US graduate medical education) [64,95] and Verona Coding Definitions for Emotional Sequences (coding of patients' emotional sequences and health professionals' responses) [94,96].

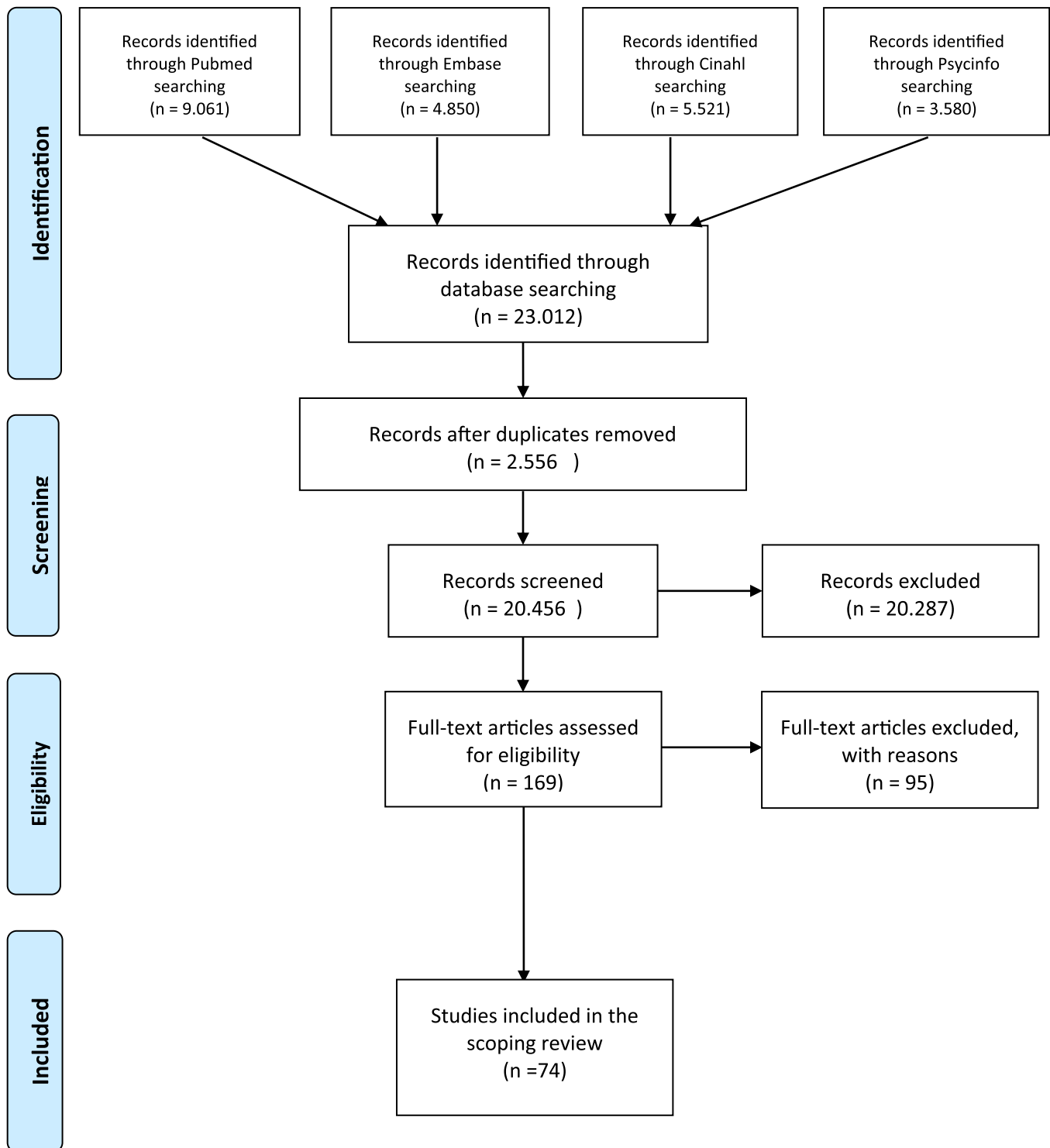


Fig. 1. Prisma written assessment flow diagram.

Among previously existing instruments ($n = 18$), the Helpful Response Questionnaire [108] in the field of motivational interviewing was the most frequently used [77,98,100,102].

3.2.4. Communication topic being assessed

Most of the studies specified which aspect of clinical communication was assessed. However, the topic was not explicitly mentioned or not clear in five articles [65,109–112]. Instruments generally focused on one topic, specific challenges were assessed more frequently than general communication skills or communication related to specific

patient groups or disciplines (Table 2). Instruments used in TDS tended to focus more often on general or basic communication exclusively ($n = 11/27$ 41%) whereas instruments used in IS tended to focus more on specific communication challenges ($n = 19/43$ 44%) (e.g. motivational interviewing) (Appendix C).

3.2.5. Type of knowledge assessed

Few studies explicitly described the type of knowledge or cognitive processes that were assessed. If explicitly mentioned, authors referred to either the Miller pyramid of clinical competence [39,69] or to

Table 1
study and study participants' characteristics (n = 74).

Study aims	Total N = 74	Designs				
Intervention studies (IS) n = 46	8	One measurement point				
	2	Cross sectional				
	2	Post one group only				
	4	Post with non-randomized control group				
	36	Two measurement points				
	25	Pre post group only				
	6	Pre post with randomized control group				
	4	Pre post with non-randomized group				
	1	Pre post more than 1 institution				
	1	None				
	1	Study protocol				
	Test development studies (TDS) n = 28	17	One measurement point			
			Most studies described test development, assessment of actual knowledge, correlation and/or validation studies with one group only			
9		Two measurement points				
		Most studies described assessment of actual knowledge, correlation and/or validation studies in pre-post designs				
2		None				
	1	Expert validation of a test				
	1	Test description without application to test-takers				
Study participants	Total					
Mean number of participants	mean (SD)					
	257.9	1. Intervention studies: 195.2 (322.4)				
	(442.69)	2. Test development studies: 383.4 (597.6)				
Participants' profession and level of training	Total N = 74	Before Admissions N = 2	Under- graduate N = 52	Post-graduate N = 11	Continuous N = 9	
One profession	66	1	50	11	4	
- Dentistry	3	0	3	0	0	
- Medicine	40	1	28	9	2	
- Nursing	9	0	8	0	1	
- Pharmacy	5	0	5	0	0	
- Psychology	3	0	3	0	0	
- Others	6	0	3	2	1	
More than one profession (Occupational therapy; Osteopathy; Physical therapy/ Physiotherapy; Bachelor Health Sciences; Counsellors in substance abuse; Staff of opioid treatment clinic)	8	1	2	0	5	

Bloom's taxonomy [90]. A few authors used verbs related to factual knowledge [92], applied knowledge [59,92,94,113], conceptual knowledge [70], practical knowledge [83], or understanding [67].

There were slightly more instruments reporting knowledge-based than behavior-based instructions (Table 2). Among the instruments containing knowledge-oriented instructions, ten of these were described in IS [52,58,63,75,84,99,114–117] and seven were described in TDS [55,74,88,89,92,96,113] (Table 2). Nine IS [59,61,76,95,98,102,109,118,119] and three TDS [50,77,100] mentioned behavior-oriented instructions. Among the instruments which did not provide any information about the lead-in question, only six were test development studies [64,65,72,81,106,120].

3.2.6. Scoring activities

Some type of instrument scoring was described in a third of the instruments (n = 25) in equal proportions between intervention and test development groups. Most reported only one criterion used to establish a respondent's score (n = 13). The most frequently reported methodology used for instrument scoring was the involvement of a group of authors/teachers (n = 15), the involvement of experts (n = 10) or the use of theoretical frameworks, guidelines or models (n = 11).

The feasibility of the scoring methodology or the administration of the instrument was rarely mentioned [51,57,62,76,84,86,95,97,120].

3.3. Item characteristics

3.3.1. Stimulus

Most authors used context-rich stimuli, with videos used as often as written vignettes (Fig. 2). No information was provided by 16 (37%) of the IS and by four – (18%) of the TDS. Instruments used in

TDS did not use context poor stimuli and relied mainly on videos (n = 13).

3.3.2. Response format

Instruments usually relied on the use of selected responses, and on constructed responses to a lesser extent (Fig. 2). A minority included both selected and constructed responses [59,64,73,90,97,102,114]. No mention of the format response was more common in the IS than TDS. For examples of different stimulus and response formats, please refer to Appendix D.

3.4. Psychometric properties

Three of the 70 studies reviewed did not report any empirical data [63–65]. Forty six studies provided information on the item and test properties, the reliability and/or validity (n = 21 for TDS and n = 25 for IS).

3.4.1. Instrument or item properties

Mean test scores were reported for 35 instruments. In one article, students' pretest and posttest scores were categorized as non-mastery (< 80% correct) or mastery (80% correct or more) instead of reporting the mean scores [84]. Standard deviation scores or comparable were reported in 27 articles. The number of respondents who passed the test, was reported by four studies [39,54,55,58] and the number of respondents who scored on different grades was reported by one [58]. Item difficulty was reported in nine studies [39,54,58,72,83,89,92,95,96]. Five studies reported the item-test correlation [72,73,80,92,100].

Table 2
Test characteristics (n = 70).

Test characteristics	N	Description
Composition of the test		
- Stand alone	19	
- Integrated with other measures	51	
	14	- Other domains such as: professionalism, clinical knowledge, etc...
	13	- Attitudes: self, efficacy, motivation, self-confidence, willingness, self-perceived competence
	6	- Behaviour/performance
	4	- Other domains and attitudes
	5	- Other domains and behaviour
	8	- Attitudes and behaviour
	1	- Other domains and attitudes and behaviour
Test development		
- Self-developed test based on	52	
- 1–2 criteria used	25	
- > 2 criteria used	10	
- None reported	17	
- Previously existing test	18	
Communication topic assessed		
- General basic communication skills exclusively	21	
- General basic communication skills in combination with specific challenges or disciplines	11	
- Specific challenges exclusively	25	
	10	- Motivational interviewing
	3	- Breaking bad news
	3	- End of life care
	4	- Cross-cultural (including work with interpreters)
	5	- Other
	4	e.g. palliative care, family medicine
- Patient groups and medical disciplines	4	
- Written communication	3	
- Mixed (general communication skills and specific challenges and/or specific discipline)	11	
- Not mentioned	5	
Type of knowledge assessed		
- Knowledge-based instructions	17	
- Behaviour-based instructions	12	
- Both	13	
- No mention about the lead-in question	28	

3.4.2. Reliability

3.4.2.1. Instrument reliability. Information about instrument reliability was provided in less than a third of the instruments, either for the whole instrument, for different versions of an instrument, for subscales of the instrument or for pre-post-interventional data. Cronbach's alpha was most commonly reported. Cronbach's alpha values ranged from .29 [55] to .92 [72]. However, not all of them were related to the assessment of communication skills (see 3.2.1). Other measures of reliability included bivariate correlation coefficient between different parts of an instrument; test-retest reliability [100]; bivariate Pearson's correlation coefficient between items related to one vignette [92]; comparison of internal consistency of a short and long test version [72]; and Kuder Richardson 20 (KR20) [72]. Two author groups used Generalisability Theory to report reliability [54,66].

3.4.2.2. Rater agreement and reliability. Less than a quarter of studies/tests reported interrater reliability. Depending on the number of raters, either Cohen's kappa or intraclass correlation was reported. Kappa and ICC values ranged between .14 [94] and .96 [100,104].

3.4.3. Validity

Information about construct validity, concurrent or predictive validity was reported for 18 instruments. In twelve studies, different approaches were used to evaluate construct validity; factor analysis [50,67]; structural equation modeling [67]; expert validation of results [74]; hypothesis driven subgroup analysis to estimate discriminatory power e.g. phases of studies, gender [50,92,96,113]; sensitivity of training effect in a pre-post design [74,75,77,84,100];

comparing an intervention and control group [92,95]; following students over time [77]; or a Content Validity Rate and Content Validity Ratio [106]. Data about concurrent validity was reported for ten instruments. These were estimated using correlation with other knowledge test results [77,88,100], correlation with measurements of behavior [68,88,92,100], correlation between test and attitudes, confidence or comparable constructs [84,89,96,104], and correlation with data from peer and patient evaluation [73]. Five studies provided information about the predictive validity of knowledge tests for later behavior [67,77,88,100,113], and one study reported incremental validity for a Situational Judgement Test [88].

3.5. Instrument availability

Only nine complete instruments were available in the article, in the appendix or online [58,59,61,76,82,100,114,116,118]. The most common forms of instrument description were a narrative description of the instrument in the text or parts/examples of the instrument given in the text or in tables. For five instruments, authors suggested that they could be contacted directly to request copies of the instrument [92,97,117,121,122].

4. Discussion and conclusion

4.1. Discussion

Written assessments of the cognitive component of communication skills are reported more frequently than we expected despite the fact that OSCEs and work-place based assessments are considered to be the best assessment measures of communication

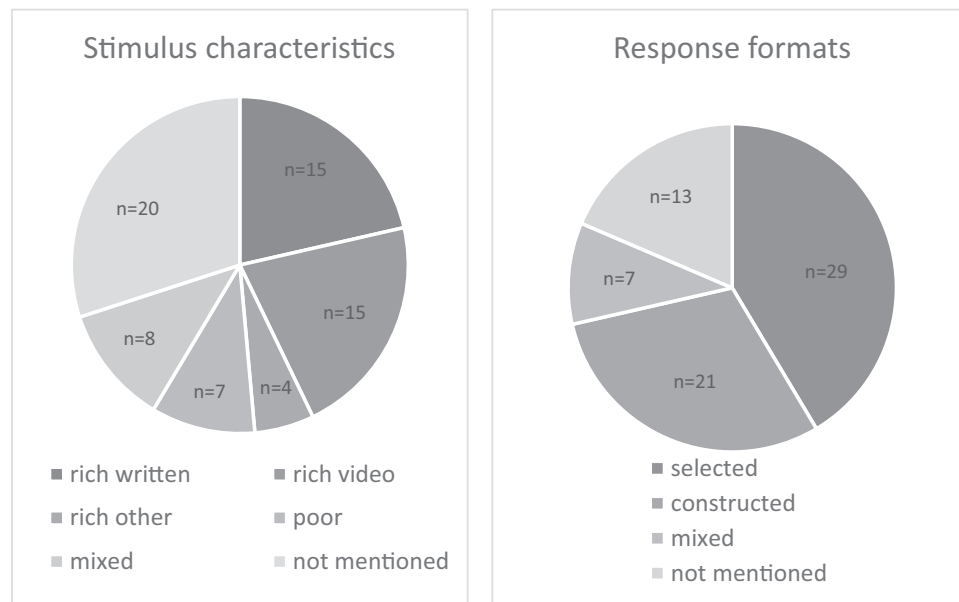


Fig. 2. Stimulus and response formats.

skills. The number of studies reporting the written assessment of the knowledge component of communication skills is increasing. Examinees at all levels of training and among a broad variety of health professionals, especially in Northern America are included. It is used for the assessment of general communication skills as well as for specific communication challenges.

The second main finding of this review is the heterogeneous quality in the reporting of the information regarding the instrument development, content, administration, feasibility, and psychometrics. Although instruments were generally developed by the authors, third of authors did not report any information about the quality criteria used in their development process. Reference to an existing theoretical framework or specific communication principles for the training or instrument development was often lacking. Psychometric properties were reported for only a minority of instruments. Both educators and researchers need to be able to evaluate whether the outcomes or scores resulting from the use of an assessment instrument can be reproduced and meaningfully interpreted in order to determine if the results are generalizable or limited to a specific context [44]. There is also a need for more information about the evidence of validity. Downing suggests five sources of validity that should be considered when interpreting scores [45]: 1) the content which refers to the match between the content of the test and the domain taught 2) the response process which covers all processes aiming at reducing the sources of error in relation to the test administration (validation of preliminary scores, quality control/accuracy of final scores, understandable description of scores for students, etc.) 3) the internal structure such as the item analysis, score scale reliability and standard errors of measurement 4) relationship to other variables –correlation with other measures of the same achievement or ability 5) consequences of the assessment on examinees, faculty, patients or society. Lack of reporting on validity estimates may be due to limited space provided by journals. It may also reflect the authors' lack of awareness about the importance of reporting the psychometric properties of an instrument, of constructive alignment between the training/topic taught and the assessment instrument [123], as well as the use of several quality control processes in order to ensure the quality of the instrument.

A third finding is that most authors did not specify the level of knowledge that the instrument was measuring with regard to clinical communication. This is surprising since a basic assessment

principle is that the choice of the question should match the level of knowledge to be measured. Several models have been developed to assist with the categorization and assessment of different levels of cognition such as Miller's pyramid or Bloom's taxonomy. The frequent use of context rich stimuli suggests that most authors assumed that the assessment of communication skills cannot be based on factual knowledge alone and that it requires some form of contextualisation. Indeed, several studies have highlighted the fact that communication challenges are context specific and that communication styles and strategies need to be adapted to the specific circumstances and goals of the consultation [124–126]. The type of instructions used in the lead-in questions also indicates that written tests of communication skills were used to assess higher levels of cognition including analysis, interpretation, creation as well as situational awareness and not just recall or understanding of factual knowledge. In this regard, Situational Judgement Tests, which were used in four of the included studies, are a promising format for assessing competencies used in ill-defined situations such as communication [55,88,92,96].

Finally, we found that intervention studies mainly used a pre-post single group design. Although analysis of the research methodology about use of written assessments of communication skills was beyond the scope of the review, we encourage researchers using written assessments to measure the impact of a training intervention to consider study designs which will result in stronger levels of evidence.

There are several limitations to this review. Written assessment of communication skills was a broad construct which includes many disciplines, approaches, definitions and topics. We found that the process of indexing relevant studies was a difficult task ending up with more than 20,000 abstracts. Despite the fact that we combined four categories of key words in order to include all relevant articles, it is possible that we missed some important studies resulting from a lack of specific indexing in several databases. We did not include studies in languages not spoken by any of the authors and did not look for the grey literature, included PhD theses not published in peer reviewed journals. The data extraction reliability was not measured in a blinded way with regard to clinical communication and two researchers were co-authors of four of the articles (CK and NJP). However, the studies were reviewed several times, first in four groups of pairs, then by a group of four authors not involved in the

Table 3
Recommendations for researchers and teachers regarding written assessment of communication skills based on a literature review [26,34,35,44,45].

Dimensions	Recommendations for researchers on how to report about written assessment of communication	Issues to consider for teachers when choosing a written assessment instrument about communication
Origin of the test	Describe if it was self-developed, based on a previously existing test or was modified according to a previously existing test	Consider for whom and in which context the test was developed in order to evaluate its transferability to your own context
Purpose of the test	Specify whether it was summative (specify if high stakes), formative or other	Define the purpose of your assessment to help you define the weight to attribute to several elements such as the use of a blueprint, the number of questions, the process of writing questions and scoring answers and the psychometric properties of the test
Topic covered (and the model/principles used if published)	Specify the topic, mention explicitly whether the teaching/assessment referred to a model, a theory and/or best practices (constructive alignment)	Foster a test that was based on a model, a theoretical framework or best practices that is in line with the content of your teaching activities
Type of knowledge measured	Refer to a taxonomy and if not possible, specify which level of cognition was assessed in the test and whether it matched the learning objectives of the activity (constructive alignment)	Ensure that the level of cognition the test assesses reflects the learning objectives of your activity. Define whether you want the questions to be knowledge or behavior-based.
Writing of questions	Mention who were the writers, to which extent they were qualified as content experts, if not, whether the questions were reviewed by experts, write whether there was a review process for clarity and sensitiveness	Pay attention to the review process used in the writing of questions to evaluate whether it respected some standards of quality
Number of items	Mention the number of questions and on which basis the authors considered that the number was sufficient to sample the content	Make sure that the number of questions included in the test is sufficient to cover the content of what has been taught, especially for summative assessment.
Type(s) of stimulus and leading question(s)	Describe the number and type of context rich (written vignette, video, transcript, etc...) and context poor stimuli	Favor context rich stimuli that refer to authentic and relevant patients' problems and to contexts closely resembling clinical practice Script concordance test, situation judgment test and reflective portfolio assessment represent high order written assessment of communication
Response format(s)	Mention the number and type of responses: – Selected response: true/false, single best answer, etc...– Constructed response: short answer, essay, reflective writing, etc...	Chose the response format according to the level of cognition you want to measure and the time and human resources you have. However, it is the content of the lead-in question or stimulus more than the response format that determines what the question tests
Type(s) of scoring	Explain how the correct/best answers were defined (theory-principles, expert-based, etc...) and what type of credits were given for the correct answers. For constructed response, provide information about how the scoring grid was established. Furthermore, address the standard setting procedure if whole tests were used for summative assessment purposes.	Pay attention to the process used in defining the scoring and evaluate whether it respected standards of quality
Psychometrics	Report the – test/item properties (means, standard deviation, item difficulty, item-test correlations)– Reliability (consistency over items, consistency over raters)– Validity (content, correlation with other variables, consequences)	Make sure that the test has strong psychometric properties if you intend to use it for summative/high stakes exams. A pass/fail cut-off point should be then defined, using an established standard setting procedure
Feasibility of scoring and administrating	Mention whether the test scoring and the test administration were feasible, what improvements should be further implemented	Assess whether the test administration and scoring is sufficiently detailed to be replicated and feasible in your own context in terms of time and human resources
Test availability	Whenever possible, make the whole test available in the article/in appendix/online	Do not hesitate to write to the authors to obtain the whole test. If necessary ask for the permission to translate the test to your own language.
Educational impact	Mention whether performing the assessment resulted in an educational effect	Check whether the test had a reported impact on future learning, OSCE scores or on workplace performance.

studies. CK and NJP scanned all studies several times to check ambiguity before finalizing the final table summarizing the main data extracted (Appendix B). Disagreements or ambiguity were solved by discussion. Finally, we did not assess the quality of study designs and results as that was beyond the scope of this review. However, we plan to further analyze the quality of data in those studies which reported information about reliability and validity.

The quality of reporting regarding the instrument development and scoring needs to be improved to allow the replication and interpretation of written assessment results and scores in the field of communication. We recommend that the questions reflect what has been taught (adequate blueprinting) [34], are in sufficient numbers to cover the domain assessed, and are clearly written and reviewed by content experts. The appropriate selection of question writers is essential [37]. The choice of the question needs to match the level of knowledge to be measured. Pretest piloting is recommended to identify flaws or lack of clarity [34]. Marking schemes need to be developed and weighted to reflect the relative importance of each

item [37]. **Otherwise, medical educators and teachers will not be able to understand when and how to use written assessment until reporting of such elements improve.**

The authors of this review are members of the assessment subgroup of the teaching committee of EACH (tEACH). We would like to provide recommendations **for researchers and teachers** based on key principles of written assessment and on our experience as clinical teachers and assessors of communication skills. These recommendations are summarized in Table 3. We see these as a starting point for further discussion and invite all researchers and teachers interested in the topic to provide feedback to us to improve these recommendations for our future work.

4.2. Conclusion

Both educators and researchers need to be aware of what and how to report information about test development regarding written assessments of communication skills in order to allow appropriate

interpretation of written assessment results and scores. Further research is needed to better delineate the advantages of written assessment in comparison with other formats. We also need to define for which group of learners written assessment is the most appropriate format. It would be of interest to evaluate whether performance on written assessments of communication early in the curriculum is correlated to clinical performance in the OSCEs or the work-place. It would allow early identification of learners who may be at risk for poor performance in the future and so facilitate the implementation of remediation for those learners.

4.3. Practical implications

Written assessment is widely used in the field of healthcare communication. Improvements in instrument development, instrument items and psychometrics reporting may help **communication skills teachers** better define when and for whom written assessment of communication is indicated. Further research is indicated to further validate the existing instruments.

CRedit authorship contribution statement

Noelle Junod Perron: Conceptualization, Methodology, Formal analysis, Writing – original draft. **Peter Pype:** Conceptualization, Methodology, Formal analysis, Writing – review & editing. **Marc van Nuland:** Conceptualization, Methodology, Formal analysis, Writing – review & editing. **Maria Magdalena Bujnowska-Fedak:**

Appendix A. : Search strategy for Pubmed

Pubmed	<p>Communication [MeSH] OR communication [Title/Abstract] OR Interpersonal Relations [Mesh:NoExp] OR Interpersonal Relation* [Title/Abstract] OR Professional-Patient Relations [Mesh] OR Professional-Patient Relation* [Title/Abstract] OR Nurse-Patient Relations [MeSH] OR Nurse-Patient Relations [Title/Abstract] OR Physician-Patient Relations [MeSH] OR Physician-Patient Relation* [Title/Abstract] OR Physician Patient Relation* [Title/Abstract] OR Doctor-Patient Relation* [Title/Abstract] OR Doctor Patient Relation* [Title/Abstract] OR Social Skills [MeSH] OR Social Skills [Title/Abstract] OR Interpersonal Skills [Title/Abstract] OR Social Competence [Title/Abstract] OR Patient-Centered Care [MeSH] OR Patient-Centered Care [Title/Abstract]</p> <p>AND</p> <p>Educational Measurement [MeSH] OR Educational Measurement [Title/Abstract] OR Educational assessment* [Title/Abstract] OR Educational Examination* [Title/Abstract] OR Written assessment* [Title/Abstract] OR Written exam* [Title/Abstract] OR Written measure [Title/Abstract] OR Written measures [Title/Abstract] OR Written measurement [Title/Abstract] OR Written measurements [Title/Abstract] OR Written test* [Title/Abstract] OR exam [Title/Abstract] OR exams [Title/Abstract] OR examination [Title/Abstract] OR examinations [Title/Abstract] OR Assessment [Title/Abstract] OR Assessments [Title/Abstract] OR Assessing [Title/Abstract] OR Assess [Title/Abstract] OR Test [Title/Abstract] OR Tests [Title/Abstract] OR Testing [Title/Abstract]</p> <p>AND</p> <p>Computer Simulation [MeSH] OR Computers [MeSH] OR Writing [MeSH] OR computer-based [Title/Abstract] OR Computer [Title/Abstract] OR computers [Title/Abstract] OR Videotape Recording [MeSH] OR Videotape Recording [Title/Abstract] OR Videotape Recordings [Title/Abstract] OR Video-based test [Title/Abstract] OR Video-based measurement [Title/Abstract] OR Video-based measurements [Title/Abstract] OR video [Title/Abstract] OR Question [Title/Abstract] OR Questions [Title/Abstract] OR Multiple choice [Title/Abstract] OR Open-ended [Title/Abstract] OR Close-ended [Title/Abstract] OR Essay [Title/Abstract] OR Script [Title/Abstract] OR Vignette[Title/Abstract] OR Situational judgement test [Title/Abstract] OR Reflective writing [Title/Abstract] OR Scenario [Title/Abstract] OR paper-and-pencil [Title/Abstract]</p> <p>AND</p> <p>Knowledge [MeSH] OR Cognition [MeSH] OR Clinical Competence [MeSH] OR Knowledge [Title/Abstract] OR Cognition [Title/Abstract] OR Clinical Competence [Title/Abstract] OR Skill [Title/Abstract] OR Skills [Title/Abstract] OR comprehension [Title/Abstract] OR cognitive process [Title/Abstract] OR cognitive process [Title/Abstract] OR analyze [Title/Abstract] OR explain [Title/Abstract] OR understand [Title/Abstract] OR evaluate[Title/Abstract]</p>
--------	---

Appendix B. : Differences between behavioural and knowledge-oriented instructions

Behavioural-oriented questions	<p>Prototypical instructions were:</p> <ul style="list-style-type: none"> • How would you proceed with the consultation? • Write here what you would say next • How would you react? (select the best from a list of given reactions)
Knowledge-oriented instructions	<p>Prototypical instructions were:</p> <hr/> <ul style="list-style-type: none"> • Identify verbal and nonverbal communication strategies (using three scenarios) • What is the best course of action for communicating with Mr. Smith? • Indicate anything the physician omitted to do or could have done better. • What is the best way to explore a patient perspective? • Which of the following strategies is most appropriate for a first step to start a conversation? • What is the most effective response? • Which of the following questions is an open question? • Where would you use a summary? • What would be an adequate next step of the medical student in the given scenario? • What procedure ensures that you communicate the information in a way that is understandable and tailored to the patient?

Conceptualization, Methodology, Formal analysis, Writing – review & editing. **Marcela Dohms:** Methodology, Formal analysis, Writing – review & editing. **Geurt Essers:** Conceptualization, Methodology, Formal analysis, Writing – review & editing. **Ragnar Joakimsen:** Conceptualization, Methodology, Formal analysis, writing - Writing – review & editing. **Zoi Tsimtsiou:** Conceptualization, Methodology, Formal analysis, Writing – review & editing. **Claudia Kiessling:** Conceptualization, Methodology, Formal analysis, Writing – review & editing.

Funding

None.

Declaration of Competing Interest

None.

Acknowledgement

We gratefully thank the librarians from the Geneva, Ghent and Tromsø Universities for their help in the literature search and Lambert Schuwirth for his constructive remarks. We also thank Eva Doherty for her feedback and for thoroughly proofreading the manuscript.

Appendix C

Articles included in the scoping review and their main characteristics regarding study aims, design, content, question and answer formats and psychometrics.

Appendix D. Examples of stimuli and responses (modified examples extracted from the articles)

Type stimulus	Type response	Examples (adapted from [61,96,108,113])
Poor stimulus	Selected response	Towards the end of your consultation, if you ask the patient an open-ended question like 'anything else you would like to ask me?' this is likely to lead to the patient with cognitive impairment: being silent being confused about what to say and not bringing up any healthcare concerns inviting to share their worries or questions with you To help a worried patient to further describe his/her complaints, certain communication skills are more adequate than others. Assuming the following interventions are given in an understanding tone of voice, in the majority of cases: 1. In reality it is often not so bad is better than it worries you, doesn't it? True/False? 2. What is it that worries you? Is better than: I have experienced something similar True/False?
Poor stimulus	Constructed response	No example found in the articles. However, here are some examples: Please explain which strategy is related to each letter of SPIKES Define what is empathy and how it helps building up the relationship with the patient (10 lines)
Rich stimulus	Selected response	Please watch the following video: https://: Please focus on the end of the film. The following questions refer to it. Task: what would be an adequate next step of the doctor in the conversation with Mrs Smith? Answer: Please click on the slider scale to active it and then drag the slider to the desired position (a line from very appropriate to very inappropriate) a) You know that such a high cholesterol level is really dangerous for you and you need to take medication. Can I explain to you the effect of the medication? b) How do you know that you will not tolerate the medication? c) Silence... d) Did you mention the side-effects of the medication to your general practitioner? e) You seem worried, Mrs Smith
Rich stimulus	Constructed response	A 47-year-old unemployed bank employee tells you: "My life has just no meaning anymore. I'm worthless. I can't find a job. Everything I try to do fails. Sometimes I wonder whether it's worth it." Write here what you would say next."

References

- [1] Spitzberg BH. (Re)Introducing communication competence to the health professions. *J Public Health Res* 2013;2:23.
- [2] Berkhof M, van Rijssen HJ, Schellart AJ, Anema JR, van der Beek AJ. Effective training strategies for teaching communication skills to physicians: an overview of systematic reviews. *Patient Educ Couns* 2011;84:152–62.
- [3] Gilligan C, Powell M, Lynagh MC, Ward BM, Lonsdale C, Harvey P, et al. Interventions for improving medical students' interpersonal communication in medical consultations. *Cochrane Database Syst Rev* 2021;2:012418.
- [4] Kerr D, Ostaszkiwicz J, Dunning T, Martin P. The effectiveness of training interventions on nurses' communication skills: a systematic review. *Nurse Educ Today* 2020;89:104405.
- [5] Kerr A, Kelleher C, Pawlikowska T, Strawbridge J. How can pharmacists develop patient-pharmacist communication skills? A realist synthesis. *Patient Educ Couns* 2021.
- [6] Hatfield TG, Withers TM, Greaves CJ. Systematic review of the effect of training interventions on the skills of health professionals in promoting health behaviour, with meta-analysis of subsequent effects on patient health behaviours. *BMC Health Serv Res* 2020;20:593.
- [7] Moore PM, Rivera Mercado S, Grez Artigues M, Lawrie TA. Communication skills training for healthcare professionals working with people who have cancer. *Cochrane Database Syst Rev* 2013;003751.
- [8] Whitehead L, Walker DK. Communication skills training for health care professionals who work with cancer patients. *Am J Nurs* 2019;119:45.
- [9] Steed L, Sohanpal R, Todd A, Madurasinghe VW, Rivas C, Edwards EA, et al. Community pharmacy interventions for health promotion: effects on professional practice and health outcomes. *Cochrane Database Syst Rev* 2019;12:011207.
- [10] Lord L, Clark-Carter D, Grove A. The effectiveness of communication-skills training interventions in end-of-life noncancer care in acute hospital-based services: a systematic review. *Palliat Support Care* 2016;14:433–44.
- [11] Chartered Society of Physiotherapy. Q.A. standards. Section 7: communication. Secondary Q.A. standards. Section 7: communication. <http://www.csp.org.uk/professional-union/professionalism/csp-expectations-members/quality-assurance-standards/section-7-co>; 2015 (Accessed 10 January 2021).
- [12] ADEA Foundation Knowledge and Skills for the New General Dentist: (As approved by the 2011 ADEA House of Delegates). *J Dent Educ* 2017;81:848–52.
- [13] National Health Service (UK). End of Life Care Strategy. Promoting High Quality Care for All Adults at the End of Life. NHS. Department of Health; 2008.
- [14] Australian Commission on Safety And Quality In Health Care. National Safety and Quality Health Service Standards. <http://nationalstandards.safetyandquality.gov.au/>; 2017 (Accessed 9 July 2021).
- [15] General Medical Council. Outcomes for graduates. 2020; <https://www.gmc-uk.org/education/standards-guidance-and-curricula/standards-and-outcomes/outcomes-for-graduates>; 2020 (Accessed 30 December 2020).
- [16] Frank JR, Snell L, Sherbino J, editors. The Draft CanMEDS 2015 Physician Competency Framework Series IV. The Royal College of Physicians and Surgeons of Canada, Ottawa: 2014.
- [17] Agency for Healthcare Research and Quality. Communicating to Improve Quality: Implementation Handbook. https://www.ahrq.gov/sites/default/files/wysiwyg/professionals/systems/hospital/engagingfamilies/strategy2/Strat2_Implement_Hndbook_508.pdf (Accessed 9 July 2021).
- [18] Swing SR. The ACGME outcome project: retrospective and prospective. *Med Teach* 2007;29:648–54.
- [19] Bachmann C, Abramovitch H, Barbu CG, Cavaco AM, Elorza RD, Haak R, et al. A European consensus on learning objectives for a core communication curriculum in health care professions. *Patient Educ Couns* 2013;93:18–26.
- [20] Dittton-Phare P, Loughland C, Duvivier R, Kelly B. Communication skills in the training of psychiatrists: a systematic review of current approaches. *Aust N Z J Psychiatry* 2017;51:675–92.
- [21] MacDonald-Wicks L, Levett-Jones T. Effective teaching of communication to health professional undergraduate and postgraduate students: a Systematic Review. *JBI Libr Syst Rev* 2012;10:1–12.
- [22] Medizinischer Fakultätentag der Bundesrepublik Deutschland (Ed.). Nationaler Kompetenzbasierter Lernzielkatalog Zahnmedizin. <http://www.nklz.de>; 2015 (Accessed 10 January 2021).
- [23] Stiefel F, Barth J, Bensing J, Fallowfield L, Jost L, Razavi D, et al. Communication skills training in oncology: a position paper based on a consensus meeting among European experts in 2009 participants. *Communication skills training in oncology: a position paper based on a consensus meeting among European experts in 2009* *Ann Oncol* 2010;21:204–7.
- [24] Norcini J, Anderson B, Bollela V, Burch V, Costa MJ, Duvivier R, et al. Criteria for good assessment: consensus statement and recommendations from the Ottawa 2010 Conference. *Med Teach* 2011;33:206–14.
- [25] Schuwirth LW, van der Vleuten CP. The use of clinical simulations in assessment. *Med Educ* 2003;37(Suppl 1):65–71.
- [26] van der Vleuten CP, Schuwirth LW, Scheele F, Driessen EW, Hodges B. The assessment of professional competence: building blocks for theory development. *Best Pr Res Clin Obstet Gynaecol* 2010;24:703–19.

- [27] Miller GE. The assessment of clinical skills/competence/performance. *Acad Med* 1990;65:S63–7.
- [28] Comert M, Zill JM, Christalle E, Dirmaier J, Harter M, Scholl I. Assessing communication skills of medical students in objective structured clinical examinations (OSCE)—a systematic review of rating scales. *PLoS One* 2016;11:0152717.
- [29] Hodges B, Turnbull J, Cohen R, Bienenstock A, Norman G. Evaluating communication skills in the OSCE format: reliability and generalizability. *Med Educ* 1996;30:38–43.
- [30] Setyungroho W, Kennedy KM, Kropmans TJ. Reliability and validity of OSCE checklists used to assess the communication skills of undergraduate medical students: A systematic review. *Patient Educ Couns* 2015.
- [31] Turner JL, Dankoski ME. Objective structured clinical exams: a critical review. *Fam Med* 2008;40:574–8.
- [32] Laidlaw A, Hart J. Communication skills: an essential component of medical curricula. Part I: Assessment of clinical communication: AMEE Guide No. 51. *Med Teach* 2011;33:6–8.
- [33] van der Vleuten C, van den Eertwegh V, Giroldi E. Assessment of communication skills. *Patient Educ Couns* 2019;102:2110–3.
- [34] Kiessling C, Tsimsiou Z, Essers G, van Nuland M, Anvik T, Bujnowska-Fedak MM, et al. General principles to consider when designing a clinical communication assessment program. *Patient Educ Couns* 2017;100:1762–8.
- [35] Schuwirth LW, van der Vleuten CP. Different written assessment methods: what can be said about their strengths and weaknesses? *Med Educ* 2004;38:974–9.
- [36] Denniston C, Molloy E, Nestel D, Woodward-Kron R, Keating JL. Learning outcomes for communication skills across the health professions: a systematic literature review and qualitative synthesis. *BMJ Open* 2017;7:014570.
- [37] Hargie O. Skill in theory: Communication as skilled performance. In: Hargie O, editor. *The handbook of communication skills*. Routledge; 2006.
- [38] Kinderman P, Humphris G. Clinical communication skills teaching: the role of cognitive schemata. *Med Educ* 1995;29:436–42.
- [39] Hulsman RL, Mollema ED, Hoos AM, de Haes JC, Donnison-Speijer JD. Assessment of medical communication skills by computer: assessment method and student experiences. *Med Educ* 2004;38:813–24.
- [40] Bloom BS. *Taxonomy of educational objectives*. McKay: New York; 1956.
- [41] Krathwohl DRA. A revision of bloom's taxonomy: an overview. *Theory into Pract* 2002;41:2012–225.
- [42] McDaniel MA, Hartman NS, Whetzel DL, Grubb III WL. Situational judgment tests, response instructions, and validity: a meta-analysis. *Pers Psychol* 2007;60:63–91.
- [43] Schuwirth LW, van der Vleuten CP. ABC of learning and teaching in medicine: Written assessment. *BMJ* 2003;326:643–5.
- [44] Downing SM. Reliability: on the reproducibility of assessment data. *Med Educ* 2004;38:1006–12.
- [45] Downing SM. Validity: on meaningful interpretation of assessment data. *Med Educ* 2003;37:830–7.
- [46] Arksey H, O'Malley L. Scoping studies: towards a Methodological Framework. *Int J Soc Res Methodol* 2005;8:19–32.
- [47] Rayyan QCRI. <https://rayyan.qcri.org/> (Accessed 11 February 2021).
- [48] Qualtrics. <https://www.qualtrics.com/about/> (Accessed 11 February 2021).
- [49] Franco C, Franco R, Severo M, Ferreira MA. A Brief Course on Clinical Communication Skills: A Multi-Centered Study. *Acta Med Port* 2016;29:809–18.
- [50] Kuntze J, Van der Molen HT, Born MP. Progress in mastery of counseling communication skills. *Eur Psychol* 2007;12:301–13.
- [51] Yeap R, Beevi Z, Lukman H, Evaluating IMU. communication skills training programme: assessment tool development. *Med J Malays* 2008;63:244–6.
- [52] Vegni E, Martinoli M, Moja EA. Improving patient-centred medicine: a preliminary experience for teaching communication skills to Italian general practitioners. *Educ Health (Abingdon)* 2002;15:51–7.
- [53] Baerheim A, Hjortdahl P, Holen A, Anvik T, Fasmer OB, Grimstad H, et al. Curriculum factors influencing knowledge of communication skills among medical students. *BMC Med Educ* 2007;7:35.
- [54] Karabilgin OS, Vatansever K, Caliskan SA, Durak HI. Assessing medical student competency in communication in the pre-clinical phase: objective structured video exam and SP exam. *Patient Educ Couns* 2012;87:293–9.
- [55] Lievens F, Sackett PR. Video-based versus written situational judgment tests: a comparison in terms of predictive validity. *J Appl Psychol* 2006;91:1181–8.
- [56] Lievens F, Sackett PR. The validity of interpersonal skills assessment via situational judgment tests for predicting academic success and job performance. *J Appl Psychol* 2012;97:460–8.
- [57] Thai TTN, Pham TT, Nguyen KT, Nguyen PM, Derese A. Can a family medicine rotation improve medical students' knowledge, skills and attitude towards primary care in Vietnam? A pre-test-post-test comparison and qualitative survey. *Trop Med Int Health* 2020;25:264–75.
- [58] Chan CH, Chan TN, Yuen MC, Tung WK. Evaluation of a simulation-based workshop on clinical performance for emergency physicians and nurses. *World J Emerg Med* 2015;6:16–22.
- [59] Cucciare MA, Ketroser N, Wilbourne P, Midboe AM, Cronkite R, Berg-Smith SM, et al. Teaching motivational interviewing to primary care staff in the veterans health administration. *J Gen Intern Med* 2012;27:953–61.
- [60] Lunsford B, Posey L. Geriatric education utilizing a palliative care framework. *Gerontol Geriatr Educ* 2018;39:183–92.
- [61] O'Brien R, Goldberg SE. The VOICE study - A before and after study of a dementia communication skills training course. *PLoS One* 2018;13:0198567.
- [62] Wittenberg-Lyles E, Goldsmith J, Ferrell B, Burchett M. Assessment of an interprofessional online curriculum for palliative care communication training. *J Palliat Med* 2014;17:400–6.
- [63] Imms C, Chu EMY, Guinea S, Sheppard L, Froude E, Carter R, et al. Effectiveness and cost-effectiveness of embedded simulation in occupational therapy clinical practice education: study protocol for a randomised controlled trial. *Trials* 2017;18:345.
- [64] Simpson D, Helm R, Drewniak T, Ziebert MM, Brown D, Mitchell J, et al. Objective structured video examinations (OSVEs) for geriatrics education. *Gerontol Geriatr Educ* 2006;26:7–24.
- [65] Langenau E, Pugliano G, Roberts W. Competency-based classification of COMLEX-USA cognitive examination test items. *J Am Osteopath Assoc* 2011;6:396–402.
- [66] Hulsman RL, Mollema ED, Oort FJ, Hoos AM, de Haes JC. Using standardized video cases for assessment of medical communication skills: reliability of an objective structured video examination by computer. *Patient Educ Couns* 2006;60:24–31.
- [67] Humphris GM. Communication skills knowledge, understanding and OSCE performance in medical trainees: a multivariate prospective study using structural equation modelling. *Med Educ* 2002;36:842–52.
- [68] Humphris GM, Kaney S. The objective structured video exam for assessment of communication skills. *Med Educ* 2000;34:939–45.
- [69] Kuntze J, van der Molen HT, Born MP. Increase in counselling communication skills after basic and advanced microskills training. *Br J Educ Psychol* 2009;79:175–88.
- [70] Lukman H, Beevi Z, Yeap R. Training future doctors to be patient-centred: efficacy of a communication skills training (CST) programme in a Malaysian medical institution. *Med J Malays* 2009;64:51–5.
- [71] Glover TL, Garvan C, Nealis RM, Citty SW, Derrico DJ. Improving end-of-life care knowledge among senior baccalaureate nursing students. *Am J Hosp Palliat Care* 2017;34:938–45.
- [72] Lange JW, Shea J, Grossman SC, Wallace M, Ferrel BR. Validation of the end-of-life nursing education consortium knowledge assessment test. *J Hosp Palliat Nurs* 2009;11:284–90.
- [73] Kim S, Brock DM, Hess BJ, Holmboe ES, Gallagher TH, Lipner RS, et al. The feasibility of a multi-format Web-based assessment of physicians' communication skills. *Patient Educ Couns* 2011;84:359–67.
- [74] Smit GN, Van Der Molen HT. Development and evaluation of a video test for the assessment of interviewing skills. *J Cancer Educ* 1995;10:195–9.
- [75] Baribeau DA, Mukovozov I, Sabljic T, Eva KW, deLottinville CB. Using an objective structured video exam to identify differential understanding of aspects of communication skills. *Med Teach* 2012;34:e242–50.
- [76] Winefield HR, Chur-Hansen A. Evaluating the outcome of communication skill teaching for entry-level medical students: does knowledge of empathy increase? *Med Educ* 2000;34:90–4.
- [77] Smith JL, Bertone P, Carpenter KM, Wain RM, Hu MC, Amrhein PC, et al. Assessment of the relationship between a written measure of empathy and an independently rated interview of Motivational Interviewing. *J Subst Abuse Treat* 2018;87:56–63.
- [78] Becker KL, Rose LE, Berg JB, Park H, Shatzer JH. The teaching effectiveness of standardized patients. *J Nurs Educ* 2006;45:103–11.
- [79] Christner JG, Stansfield RB, Schiller JH, Madenci A, Keefer PM, Pituch K. Use of simulated electronic mail (e-mail) to assess medical student knowledge, professionalism, and communication skills. *Acad Med* 2010;85:S1–4.
- [80] Ferrell BR, Virani R, Grant M, Rhome A, Malloy P, Bednash G, et al. Evaluation of the End-of-Life Nursing Education Consortium undergraduate faculty training program. *J Palliat Med* 2005;8:107–14.
- [81] Klein D, Nagji A. Assessment of communication skills in family medicine. *Can Fam Physician* 2015;61:e412–6.
- [82] Bittner A, Jonietz A, Bittner J, Beickert L, Harendza S. Translating medical documents into plain language enhances communication skills in medical students—A pilot study. *Patient Educ Couns* 2015;98:1137–41.
- [83] Fischbeck S, Deister T, Schneider S. Teaching patient-oriented communication skills: pilot study for a development and evaluation of a practical examination in preclinical medical studies. *Z Med Psychol* 2009;5–12.
- [84] Poirier MK, Clark MM, Cerhan JH, Pruthi S, Geda YE, Dale LC. Teaching motivational interviewing to first-year medical students to improve counseling skills in health behavior change. *Mayo Clin Proc* 2004;79:327–31.
- [85] Smith RC, Lyles JS, Mettler J, Stoffelmayr BE, Van Egeren LF, Marshall AA, et al. The effectiveness of intensive training for residents in interviewing. A randomized, controlled study. *Ann Intern Med* 1998;128:118–26.
- [86] Choi H, Lee U, Jeon YS, Kim C. Efficacy of the computer simulation-based, interactive communication education program for nursing students. *Nurse Educ Today* 2020;91:104467.
- [87] Crawford P, Aubeeluck A, Brown B, Cotel-Gibbons L, Porock D, Baker C. An evaluation of a DVD trigger based assessment of communication and care delivery skills. *Nurse Educ Today* 2009;29:456–63.
- [88] Cullen MJ, Zhang C. Improving our ability to predict resident applicant performance: validity evidence for a situational judgment. *Teach Learn Med* 2020;32:508–21.
- [89] Hudelson P, Perneger T, Kolly V, Perron NJ. Self-assessed competency at working with a medical interpreter is not associated with knowledge of good practice. *PLoS One* 2012;7:38973.
- [90] Kim S, Spielberg F, Mauksch L, Farber S, Duong C, Fitch W, et al. Comparing narrative and multiple-choice formats in online communication skill assessment. *Med Educ* 2009;43:533–41.

- [91] Wittenberg E, Goldsmith JV, Williams Y, Lee A. Caring for Family Caregivers: a Pilot Test of an Online COMFORT (SM) Communication Training Module for Undergraduate Nursing Students. *J Cancer Educ* 2018.
- [92] Kiessling C, Bauer J, Gartmeier M, Iblher P, Karsten G, Kiesewetter J, et al. Development and validation of a computer-based situational judgement test to assess medical students' communication skills in the field of shared decision making. *Patient Educ Couns* 2016;99:1858–64.
- [93] Kulkarni KK, Bhalearo PM, Sarode VS. Communication skills training for medical students: a prospective, observational, single centre study. *J Clin Diagn Res* 2018;12:JC14–7.
- [94] Ortwein H, Benz A, Carl P, Huwendiek S, Pander T, Kiessling C. Applying the Verona coding definitions of emotional sequences (VR-CoDES) to code medical students' written responses to written case scenarios: Some methodological and practical considerations. *Patient Educ Couns* 2017;100:305–12.
- [95] Wroe EB, McBain RK, Michaelis A, Dunbar EL, Hirschhorn LR, Cancedda C. A novel scenario-based interview tool to evaluate nontechnical skills and competencies in global health delivery. *J Grad Med Educ* 2017;9:467–72.
- [96] Graupe T, Fischer MR, Strijbos JW, Kiessling C. Development and piloting of a Situational Judgement Test for emotion-handling skills using the Verona Coding Definitions of Emotional Sequences (VR-CoDES). *Teach Learn Med* 2020.
- [97] Biddle Jr. MA, Hoover RM. Teaching motivational interviewing in a blended learning environment. *Curr Pharm Teach Learn* 2020;12:728–34.
- [98] Childers JW, Bost JE, Kraemer KL, Cluss PA, Spagnoletti CL, Gonzaga AM, et al. Giving residents tools to talk about behavior change: a motivational interviewing curriculum description and evaluation. *Patient Educ Couns* 2012;89:281–7.
- [99] Fortune J, Breckon J, Norris M, Eva G, Frater T. Motivational interviewing training for physiotherapy and occupational therapy students: effect on confidence, knowledge and skills. *Patient Educ Couns* 2019;102:694–700.
- [100] Hartzler B. Adapting the helpful responses questionnaire to assess communication skills involved in delivering contingency management: preliminary psychometrics. *J Subst Abus Treat* 2015;55:52–7.
- [101] Keifenheim KE, Velten-Schurian K, Fahse B, Erschens R, Loda T, Wiesner L, et al. "A change would do you good": training medical students in Motivational Interviewing using a blended-learning approach - A pilot evaluation. *Eur J Dent Educ* 2019;102:663–9.
- [102] Martino S, Haeseler F, Belitsky R, Pantaloni M, Fortin AHT. Teaching brief motivational interviewing to Year three medical students. *Med Educ* 2007;41:160–7.
- [103] Villaume WA, Berger BA, Barker BN. Learning motivational interviewing: scripting a virtual patient. *Am J Pharm Educ* 2006;70:33.
- [104] Goldsmith J, Wittenberg-Lyles E, Frisby BN, Platt CS. The entry-level physical therapist: a case for COMFORT communication training. *Health Commun* 2015;30:737–45.
- [105] Choudhary A, Gupta V. Teaching communications skills to medical students: Introducing the fine art of medical practice. *Int J Appl Basic Med Res* 2015;5:S41–4.
- [106] Azadi A, Abdekhoda M, Habibi S. Medical student's skills in notifying bad news based on SPIKES protocol. *J Liaquat Uni Med Health Sci* 2018;17:249–54.
- [107] Ghoneim N, Dariya V. Teaching NICU fellows how to relay difficult news using a simulation-based curriculum: does comfort lead to competence? *Teach Learn Med* 2019;31:207–21.
- [108] Miller WR, Hedrick KE, Orlofsky DR. The Helpful Responses Questionnaire: a procedure for measuring therapeutic empathy. *J Clin Psychol* 1991;47:444–8.
- [109] Aper L, Reniers J, Koole S, Valcke M, Derese A. Impact of three alternative consultation training formats on self-efficacy and consultation skills of medical students. *Med Teach* 2012;34:e500–7.
- [110] Chang WY, Hsiao Sheen ST, Chang PC, Lee PH. Developing an E-learning education programme for staff nurses: processes and outcomes. *Nurse Educ Today* 2008;28:822–8.
- [111] Kalet AL, Mukherjee D, Felix K, Steinberg SE, Nachbar M, Lee A, et al. Can a web-based curriculum improve students' knowledge of, and attitudes about, the interpreted medical interview? *J Gen Intern Med* 2005;20:929–34.
- [112] Lonie JM, Rahim H. Does the addition of writing into a pharmacy communication skills course significantly impact student communicative learning outcomes? A pilot study. *J Pharm Pr* 2010;23:525–30.
- [113] van Dalen J, Kerkhofs E, Verwijnen GM, van Knippenberg-van den Berg BW, van den Hout HA, Scherpbier AJ, et al. Predicting communication skills with a paper-and-pencil test. *Med Educ* 2002;36:148–53.
- [114] Arif S, Cryder B, Mazan J, Quinones-Boex A, Cyganska A. Using patient case video vignettes to improve students' understanding of cross-cultural communication. *Am J Pharm Educ* 2017;81:56.
- [115] Ikram UZ, Essink-Bot ML, Suurmond J. How we developed an effective e-learning module for medical students on using professional interpreters. *Med Teach* 2015;37:422–7.
- [116] Sukalich S, Elliott JO, Ruffner G. Teaching medical error disclosure to residents using patient-centered simulation training. *Acad Med* 2014;89:136–43.
- [117] van der Molen HT, Klaver AA, Duyx MP. Effectiveness of a communication skills training programme for the management of dental anxiety. *Br Dent J* 2004;196:101–7.
- [118] Oakley M, Horvath Z, Weinberg SM, Bhatt J, Spallek H. Comparison of e-mail communication skills among first- and fourth-year dental students. *J Dent Educ* 2013;77:1413–24.
- [119] Papadopoulos L, Pentzou AE, Louloudiadis K, Tsiatsos TK. Design and evaluation of a simulation for pediatric dentistry in virtual worlds. *J Med Internet Res* 2013;15:240.
- [120] Aina B, Ogunbiyi O. Assessment of communication skills among Pharmacy students of the University of Lagos, Lagos, Nigeria. *J Basic Clin Pharm* 2011;3:215–8.
- [121] Fischer GS, Arnold RM. Feasibility of a brief workshop on palliative care communication skills for medical interns. *J Palliat Med* 2007;10:19–23.
- [122] Kim HY, Ko E, Lee ES. Effects of simulation-based education on communication skill and clinical competence in maternity nursing practicum. *Korean J Women Health Nurs* 2012;18:312–20.
- [123] Cegala DJ, Lenzmeier Broz S. Physician communication skills training: a review of theoretical backgrounds, objectives and skills. *Med Educ* 2002;36:1004–16.
- [124] Cox A, Li S. The medical consultation through the lenses of language and social interaction theory. *Adv Health Sci Educ Theory Pr* 2020;25:241–57.
- [125] Essers G, van Dulmen S, van Weel C, van der Vleuten C, Kramer A. Identifying context factors explaining physician's low performance in communication assessment: an explorative study in general practice. *BMC Fam Pr* 2011;12:138.
- [126] Salmon P, Young B. Creativity in clinical communication: from communication skills to skilled communication. *Med Educ* 2011;45:217–26.