



OPEN ACCESS

EDITED BY

Sidra Ishaque,
Aga Khan University, Pakistan

REVIEWED BY

Sonia Lorente,
Consorti Sanitari de Terrassa, Spain
Takanari Ikeyama,
Aichi Child Health and Medical General
Center, Japan

*CORRESPONDENCE

Orsola Gawronski
orsola.gawronski@opbg.net

SPECIALTY SECTION

This article was submitted to
Healthcare Professions Education,
a section of the journal
Frontiers in Medicine

RECEIVED 05 July 2022

ACCEPTED 11 October 2022

PUBLISHED 03 November 2022

CITATION

Gawronski O, Thekkan KR, Genna C,
Egman S, Sansone V, Erba I, Vittori A,
Varano C, Dall'Oglio I, Tiozzo E and
Chiusolo F (2022) Instruments
to evaluate non-technical skills during
high fidelity simulation: A systematic
review.

Front. Med. 9:986296.

doi: 10.3389/fmed.2022.986296

COPYRIGHT

© 2022 Gawronski, Thekkan, Genna,
Egman, Sansone, Erba, Vittori, Varano,
Dall'Oglio, Tiozzo and Chiusolo. This is
an open-access article distributed
under the terms of the [Creative
Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).
The use, distribution or reproduction in
other forums is permitted, provided
the original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which
does not comply with these terms.

Instruments to evaluate non-technical skills during high fidelity simulation: A systematic review

Orsola Gawronski^{1*}, Kiara R. Thekkan¹, Catia Genna¹,
Sabrina Egman², Vincenza Sansone¹, Ilaria Erba³,
Alessandro Vittori³, Carmelita Varano⁴,
Immacolata Dall'Oglio¹, Emanuela Tiozzo¹ and
Fabrizio Chiusolo³

¹Professional Development, Continuing Education and Research Unit, Bambino Gesù Children's Hospital (IRCCS), Rome, Italy, ²Clinical Risk, Innovation and Integration of Care Services, Bambino Gesù Children's Hospital (IRCCS), Rome, Italy, ³Department of Anesthesia and Critical Care, Bambino Gesù Children's Hospital (IRCCS), Rome, Italy, ⁴Department of Pediatric Cardiology and Cardiac Surgery, Bambino Gesù Children's Hospital (IRCCS), Rome, Italy

Introduction: High Fidelity Simulations (HFS) are increasingly used to develop Non-Technical Skills (NTS) in healthcare providers, medical and nursing students. Instruments to measure NTS are needed to evaluate the healthcare providers' (HCPs) performance during HFS. The aim of this systematic review is to describe the domains, items, characteristics and psychometric properties of instruments devised to evaluate the NTS of HCPs during HFS.

Methods: A systematic review of the literature was performed according to the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA). Studies were retrieved from PubMed, Cinahl, Web of Science, Cochrane Library, ProQuest and PubPsych. Studies evaluating the measurement properties of instruments used to assess NTS during HFS training were included. Pairs of independent reviewers determined the eligibility, extracted and evaluated the data. Risk of bias and appraisal of the methodological quality of the studies was assessed using the Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) checklist, and the quality of the evidence with the Grading of Recommendations, Assessment, Development and Evaluation (GRADE).

Results: A total of 3,953 articles were screened. A total of 110 reports were assessed for eligibility and 26 studies were included. Studies were conducted in Europe/United Kingdom ($n = 13$; 50%), North America/Australia ($n = 12$; 46%) and Thailand ($n = 1$; 4%). The NTS instruments reported in this review included from 1 to 14 domains (median of 4, $Q_1 = 3.75$, $Q_3 = 5$) and from 3 to 63 items (median of 15, $Q_1 = 10$, $Q_3 = 19.75$). Out of 19 NTS assessment instruments for HFS, the Team Emergency Assessment Measure (TEAM) can be recommended for use to assess NTS. All the other instruments require further research to assess their quality in order to be recommended for use

during HFS training. Eight NTS instruments had a positive overall rating of their content validity with at least a moderate quality of evidence.

Conclusion: Among a large variety of published instruments, TEAM can be recommended for use to assess NTS during HFS. Evidence is still limited on essential aspects of validity and reliability of all the other NTS instruments included in this review. Further research is warranted to establish their performance in order to be reliably used for HFS.

KEYWORDS

high fidelity simulation, non-technical skills, crew resource management, teamwork, human error, psychometrics, assessment and evaluation, reproducibility of results

Introduction

Adverse events and deaths due to human error are still significant in different fields of healthcare in spite of diagnostic advancements and their therapeutic options (1). Errors during emergency situations on hospital wards have been found to be related not to medical knowledge but to the way this is applied in complex and multidisciplinary settings (2, 3). Various reports point out that human factors contribute to 43–70% of adverse events in emergency and operating room settings (4–9). Communication breakdowns, lack of leadership and teamwork, lack of knowledge of the work environment and failed closed loop communication affect the patient care process in acute and intensive care wards (10, 11).

“Crisis resource management” (CRM) is a simulation-based training program adapted from aviation to healthcare teams for teaching non-technical skills (NTS) to healthcare providers (HCPs) and optimize team performance during patient emergencies and critical events (12–14). NTS are interpersonal cognitive, social and personal management skills that are an adjunct to technical skills (TS), contributing to safe and efficient task performance (15–17). NTS involve effective teamwork, leadership, communication, decision making, situational awareness, task and resource management (12, 18–21).

Simulation training of resuscitation team members is highly recommended by the American Heart Association (AHA) and International Liaison Committee on Resuscitation (ILCOR) (22, 23). High fidelity simulation (HFS) training enables the acquisition of critical thinking, TS and NTS through experiential learning using sophisticated life-like manikins in a realistic patient environment. High fidelity simulations are characterized by a high level of realism associated with the simulation activity, including physical (environment, equipment), psychological (emotions, situational awareness) and social factors (group culture, goals and motivations) (24,

25). Simulation-based training in pediatric critical care settings has resulted in improvements in knowledge and safety attitudes by reflecting on clinical situations and early recognition or management of conditions of risk to patient safety (26–29). Simulators and audio/video-recording of simulation training enable participants and teams to improve their skills through debriefing and replay. NTS are increasingly evaluated in medical and nursing students to evaluate their competences and the effect of simulation training on the acquisition of NTS (28).

The identification of valid and reliable instruments for the evaluation of NTS provides an opportunity to standardize their evaluation in HFS programs and avoid measurement errors (30). To date, several tools have been developed for this purpose but no gold standard has been established to evaluate NTS. The aim of this review is to identify and describe the domains, items, characteristics and psychometric properties of published instruments to evaluate NTS of HCPs during HFS.

Methods

Study design

The systematic review was conducted according to the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) (31). Our review aims to answer the following research questions: “What are the characteristics of published instruments to measure team NTS during HFS in healthcare?”; “What are the measuring properties of those instruments?”; “Are the instruments valid and reliable?”

Search strategy

A systematic search was performed on the following databases: PubMed, US National Library of Medicine, by

National Center for Biotechnology Information (NCBI), CINAHL Cumulative Index to Nursing and Allied Health Literature, by EBSCOhost, Web of Science Core Collection™ by Clarivate, Cochrane Library by The Cochrane Collaboration, ProQuest by ProQuest LLC and PubPsych by applying the filter “Human” to the search, which we conducted in July 2021 with no time limits and updated in September 2022.

The key words identified and used to formulate the search strategy were: “Simulation Training,” “High Fidelity Simulation Training,” “Assessment,” “Evaluation,” “human factor*,” “resource management,” “stress management,” “resource utilization,” “task management,” “human error,” “non-technical skill*,” “nontechnical skill*,” “Intersectoral Collaboration,” “Crew Resource Management, Healthcare,” “Leadership,” “Decision Making,” “Situation awareness,” “Communication,” “Team work,” “Team-work,” and “Teamwork.” The PubMed search strategy was peer-reviewed by a PhD prepared nurse, expert in systematic reviews. The search strategy for PubMed, CINAHL, and Cochrane is reported on [Supplementary Table 1](#).

Eligibility criteria

Studies eligible for inclusion reported the characteristics, validity or reliability of NTS evaluation instruments applied to HFS training in healthcare. NTS included communication skills, leadership, teamwork, situation awareness, decision making and task management.

The inclusion criteria were the following: (1) articles describing the characteristics, the measurement properties and performance of NTS evaluation instruments applied to HFS in healthcare; (2) instruments designed for use by direct observations of or audio-visual recordings of HFS.

The exclusion criteria were: (1) different context from healthcare training; (2) low fidelity simulations defined as: simulations using role playing or task trainers designed for specific tasks or procedures for student learning “not needing to be controlled or programmed externally for the learner to participate” (25, 32); (3) no evaluation of the instruments’ validity or reliability; (4) evaluation of TS only; (5) systematic reviews; (6) unavailability of full texts; (7) language other than English.

Study screening

Six independent reviewers screened the titles and abstracts for eligibility according to inclusion criteria, followed by the full-texts of articles identified as “included” or “interesting.” Each article was evaluated on a double-blinded basis by two researchers independently. The Rayyan Intelligent Systematic Review System was used to perform blinded electronic screening (33). Preliminary training on its use was performed on a selected

database. Disagreements at each level of screening were resolved through consensus discussion or assistance by another reviewer if needed. Duplicate records were identified and removed. The search process and number of articles retrieved and excluded at each step of the process are shown in the PRISMA flow diagram ([Figure 1](#)).

Data extraction

Data were extracted from each paper included in the review by two authors independently using three tables on Microsoft Word. Extracted information was verified by a third reviewer. Data extracted included: n° of domains, items and type of NTS, scale used to describe the characteristics of NTS instruments; country, study objective, population, type of scenario, the psychometric properties assessed according to the COSMIN criteria, and the final GRADE assessment.

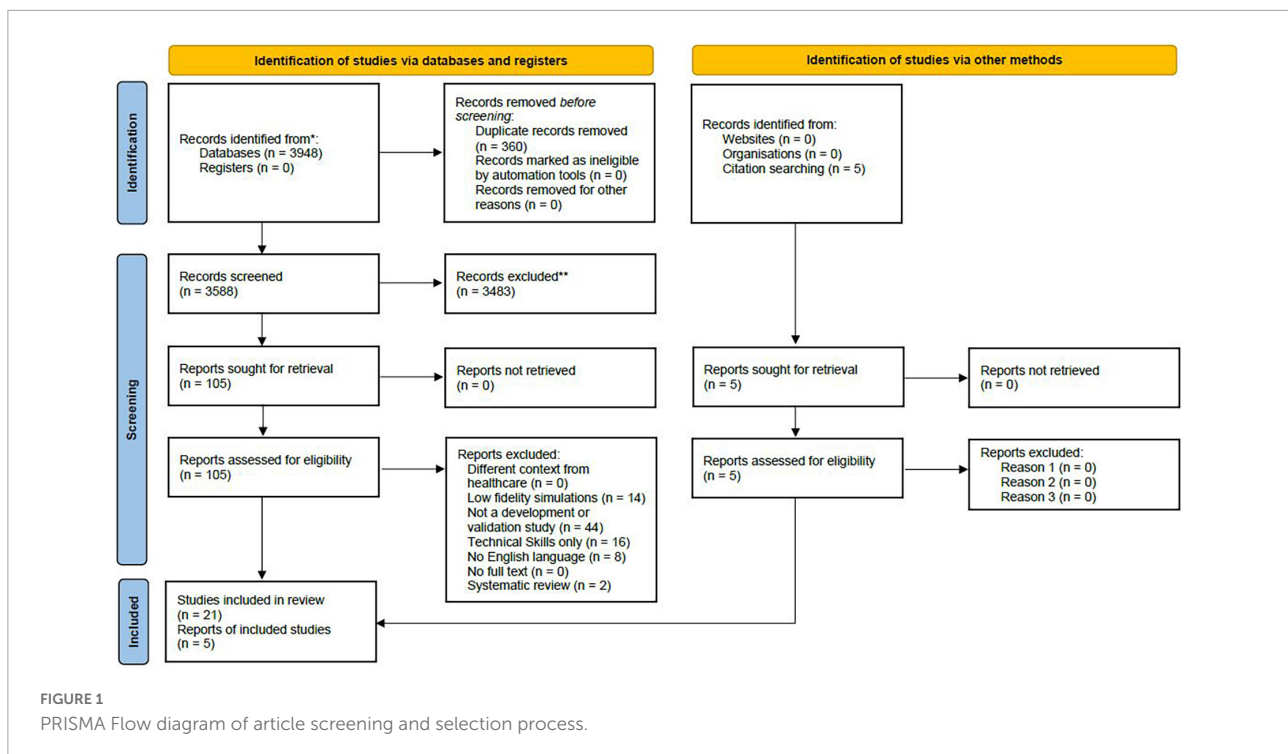
Quality assessment

Methodological quality of included studies by measurement properties

The methodological quality of the included articles was evaluated using the Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) checklist (34–36).

The COSMIN checklist is a validated and standardized quality assessment tool and is increasingly used in systematic reviews of instrument measurement properties. The COSMIN checklist provides clear evaluation criteria and standards for the methodological quality of studies that report and evaluate the psychometric properties of measurement tools. Over a total of 88 items, 70% are related to the study design phase (35). The COSMIN method requires assessing the methodological quality of each study across 10 boxes, including questions referred to tool development, content validity, structural validity, internal consistency, cross-cultural validity, reliability, measurement error, criterion validity, hypothesis testing for construct validity and responsiveness.

The research team met to review and discuss the COSMIN criteria, methods and ratings to ensure a standardized approach in accordance with the guidelines. Six research team members independently scored, in sub-teams of two, each set of questions on a 4-point rating scale (“Inadequate,” “Doubtful,” “Adequate” or “Very Good”) based on the COSMIN criteria and checklists. Standardized forms in Excel were used to report the assessments. Any disagreement in the ratings were resolved by negotiation by the two reviewers, with a third reviewer involved when necessary. In accordance with the COSMIN guidelines, methodological quality scores for each NTS instrument were assigned by taking the lowest rating of any item in each box.



Overall rating of non-technical skills instruments by measurement properties

To obtain an overall rating on the identified NTS instruments, the results obtained for each measurement property were evaluated. Based on the criteria for a good measurement property, each result obtained from each study was assessed as “sufficient,” “insufficient,” “inconsistent,” or “indeterminate.” Consistent results of more studies on the same NTS instrument were grouped together and the overall rating (OR) on measurement properties of each identified NTS instrument was determined, according to the COSMIN criteria.

Quality of evidence of non-technical skills instruments by measurement properties

Finally, we used a modified GRADE (Grading of Recommendations, Assessment, Development, and Evaluation) approach to ensure strength and certainty of evidence. The quality of evidence for each measurement property was graded as “high,” “moderate,” “low” or “very low” based on methodological quality and overall rating, according to the COSMIN manual.

Results

Study selection

Our search produced a total of 3,953 potentially eligible studies. After duplicates were removed, the titles and abstracts of

the remaining 3,593 studies were screened. A total of 110 papers were assessed for eligibility, and 26 were included in the review as relevant to the research question (Figure 1).

Study characteristics

In this review, 19 NTS assessment instruments for HFS were found. Studies were conducted in the following regions or countries: Europe/United Kingdom ($n = 13$; 50%), North America/Australia ($n = 12$; 46%) and Thailand ($n = 1$; 4%).

All studies involved simulation training with HCPs or students that practice in hospital and/or university settings. The sample size of the included studies was between 5 (37) and 177 HCPs/students (13). Four studies did not report the number of participants involved (38–41). HCPs included physicians (4 studies; 16%), midwives (1 study; 4%), medical or nursing students (9 studies; 34%), and multiprofessional teams of HCPs (12 studies; 46%). The NTS instruments were used during simulations of in-hospital patient emergencies. Simulation scenarios focused on general ($n = 9$; 36%), surgical ($n = 4$; 16%), obstetric and gynecological ($n = 4$; 16%), pediatric ($n = 3$; 12%), trauma ($n = 3$; 12%) and operating room emergencies ($n = 2$; 8%). One study did not report the type of simulation scenario. **Supplementary Table 2** reports the NTS instruments ordered by type of scenario.

The NTS instruments reported in this review included 1 to 14 domains (median of 4) and 3–63 items (median of 15). The

most frequent domains were Leadership (16; 84%), Teamwork (14; 79%), Situational awareness (13; 68%), Communication (13; 63%), Task management (9; 48%), and Decision making (9; 47%). All the instruments used a Likert scale to measure the assessment, with scores that ranged from a minimum of 0 to a maximum of 9 points (median of 1–5 Likert point scale). The characteristics of the included studies and NTS instruments are reported in [Tables 1, 2](#).

Quality assessment

Measurement properties of assessment tools

The development of the NTS instruments were reported in 14 studies. The quality of the development of the NTS instruments was doubtful or inadequate, due to missing substantial elements for an adequate development process according to the COSMIN method. The 26 included studies reported NTS instrument measurement properties, primarily on: reliability ($n = 22$; 85%), internal consistency ($n = 17$; 65%), and content validity ($n = 11$; 42%). For a total of 19 NTS assessment instruments, the following measurement properties were reported: content validity ($n = 19$), internal consistency ($n = 12$), reliability ($n = 17$), construct validity and responsiveness ($n = 6$), structural validity ($n = 4$), cross cultural validity ($n = 2$), criterion validity ($n = 1$), and measurement error ($n = 1$). The methodological quality of the included studies by measurement properties of NTS instruments is reported in [Table 2](#).

Overall ratings of the measurement properties of the NTS instruments and the quality of evidence are reported in [Table 3](#). [Figure 2](#) shows the overall methodological quality of the studies included in this review, the good measurement properties and the quality of evidence (\geq moderate) of the NTS instruments reported in this review. In four studies, the usability and feasibility of the NTS instruments Anesthetists' Non-Technical Skills (ANTS), Anaesthesiology Students' Non-Technical Skills (AS-NTS), Assessment of Obstetrical Team Performance (AOTP)/Global Assessment of Obstetrical Team Performance (GAOTP) and Team Emergency Assessment Measure (TEAM) was assessed through a survey administered to the users (17, 42–44).

Recommendations

Of the 19 instruments, according to the COSMIN criteria, the Team Emergency Assessment Measure (TEAM) can be recommended for the assessment of NTS as it has sufficient evidence of content validity and at least low quality evidence for sufficient internal consistency. TEAM was developed in Australia for trained observers to rate team performance during simulated resuscitations and deliver constructive debriefing

sessions. The instrument has 3 domains (leadership, teamwork, and task management) and 11 items. TEAM was studied in the medical emergency and obstetric/gynecologic settings. TEAM's content validity (content validity index = 0.96) and internal consistency was reported to be high (ranging from Cronbach's $\alpha = 0.85$ – 0.92) while its reliability, was moderate-to-high (ICC = 0.66–0.98) (44–46). Overall acceptability and satisfaction with the use of TEAM was high, including design and observability of the teamwork skills. In one study, the item “team morale” was rated as difficult to determine, particularly for raters without previous experience with resuscitation events (44).

All the other instruments included in this review require further research to assess their quality in order to be recommended for use during simulation training. Eight NTS instruments had a positive overall rating of their content validity with at least a moderate quality of evidence, but would need further testing to be recommended, regarding construct validity and internal consistency. None of the instruments included in this review were considered not recommendable for HFS training because there was no high-quality evidence that confirmed the inadequacy of their psychometric properties.

Discussion

This systematic review applied the COSMIN methodology to assess the psychometric properties of instruments measuring NTS. A total of 19 instruments to evaluate NTS during high fidelity simulations were identified. One instrument, TEAM, fulfilled the psychometric testing requirements for the recommendation of its use during HFS training according to the COSMIN criteria. All the other instruments require further testing as they did not report sufficient evidence of content validity or at least show low-quality evidence for sufficient internal consistency. None of the instruments were considered not recommendable, suggesting that there is margin to further investigate and report the essential measuring properties required for recommending their use.

Patient safety is receiving increasing attention in HCPs' curricula. The importance of NTS training is consistently emerging as an essential component of safety competence for HCPs (47–49). Measuring NTS during HFS training through validated instruments is essential to monitor their development and improve their awareness among medical, nursing students and HCPs (50). The effectiveness of HFS programs is based on their ability to achieve their educational goals, which may include both technical and NTS. While TS are commonly evaluated through standardized instruments, NTS are seldom evaluated (51–53). Increasing evidence of the effect of simulation

TABLE 1 Characteristics of the included non-technical skills instruments.

NTS instruments	N° of domains	N° of items	Teamwork	Leadership	Communication	Decision making	Situation awareness	Task management	Other variables	Likert scales (range)
ANTS	4	15	✓			✓	✓	✓		1–4
ANTSdk	4	16	✓	✓		✓	✓			1–5
ASNTS	3	3	✓	✓				✓	Problem solving; Team orientation	1–5
Assessment of EM physicians' NTS	4	12	✓	✓		✓	✓			1–9
AOTP-GAOTP	6	18	✓		✓		✓	✓	Environment in the room Communication with patient and partner	1–5
BARS tool	4	4	✓		✓	✓	✓			1–9
CALM	4	16	✓	✓	✓			✓		1–4
Ottawa GRS	5	6		✓	✓		✓	✓	Problem solving	1–7
HFRS	5	45	✓	✓	✓				Confidence assertion; error	1–5
LOSA	4	13		✓	✓		✓		Preoperative preparation	1–5
MHPTS	1	16	✓	✓	✓					0–2
NOTSS	4	12	✓	✓		✓	✓			1–4
NTS-NAS	14	63	✓	✓	✓	✓	✓	✓	Know the environment, Call for help, Distribute the workload, Use all information, Prevent fixation errors, Cross check, Use cognitive aids, Allocate attention	1–5
OSANTS	7	7	✓	✓	✓	✓	✓		Professionalism, Managing and Coordinating	1–5
OSCAR	6	48		✓	✓	✓	✓		Cooperation and Coordination	0–6
STAT	3	26	✓	✓				✓	68 Technical Skills items	0–2
T-NOTECHS	1	5		✓	✓	✓	✓	✓		1–5
TEAM	3	11	✓	✓				✓		0–4
TPOT	5	25	✓	✓	✓		✓		Team Structure	1–5

ANTS, Anesthetists' Non-Technical Skills; ANTSdk, Anesthesiologists' Non-Technical Skills in Denmark; AOTP, Assessment of Obstetrical Team Performance; ASNTS, Anesthesiology students' Non-Technical Skills; BARS, Behaviorally Anchored Rating Scale; CALM, Concise Assessment of Leader Management; GAOTP, Global Assessment of Obstetrical Team Performance; HFRS, Human Factors Rating Scale; MHPTS, Mayo High Performance Team Scale; NOTSS, Non-Technical Skills for Surgeons; NTS, Non-Technical skills; NTS-NAS, Non-Technical Skills-Nursing Assessment Scale; OSANTS, Objective Structured Assessment of Non-Technical Skills; OSCAR, Observational Skill-based Clinical Assessment tool for Resuscitation; Ottawa GRS, Global Rating Scale; STAT, Simulation Team Assessment Tool; T-NOTECHS, Non-technical skills scale for trauma; TEAM, Team Emergency Assessment Measure; TPOT, Team Performance Observational Tool.

TABLE 2 Methodological quality of the included studies by measurement properties of NTS instruments.

NTS instruments (references)	Population	Scenario	Instrument design	Content validity	Risk of bias								
					Structural validity	Internal consistency	Cross cultural validity	Reliability	Measurement error	Criterion validity	Construct validity	Responsiveness	
ANTS	(42)	50 anesthesiologists	Anesthesiology	I	D		D		I	V			
	(65)	70 anesthesiology R	Operating room emergency	I	D		I		V			A	A
ANTSdk	(66)	31 HCPs	General surgery	I									
	(67)	19 anesthesiologists	Anesthesiology	I					V				A
ASNTS	(17)	21 anesthesiology R	Anesthesiology/emergency	D	D				D				
Assessment of EM physicians' NTS	(55)	148 HCPs	Emergency	D	A								
AOTP-GAOTP	(43)	72 obstetrical team members	Obstetric emergencies	I	D		I		V				
BARS tool	(41)	Anesthesiology R and NS	Pediatric emergency	I					V				A
CALM	(68)	40 HCPs	Pediatric emergency	D	D				D				
GRS	(12)	59 R	Emergency	I	D		D		V				
	(64)	16 N, 6 O, 6 anesthesiologists, and 6 R	Obstetric emergencies	I			I		V				
	(69)	28 R	Medical/Obstetric Emergency	I			I	I	V				
HFRS	(65)	70 anesthesiology R	Operating room emergency	I	D		I		V			A	A
	(64)	16 N, 6 O, 6 anesthesiologists, and 6 R	Obstetric emergencies	I			I		V				

(Continued)

TABLE 2 (Continued)

NTS instruments (references)	Population	Scenario	Instrument design	Content validity	Risk of bias							
					Structural validity	Internal consistency	Cross cultural validity	Reliability	Measurement error	Criterion validity	Construct validity	Responsiveness
LOSA (70)	27 surgical R	Respiratory emergencies	D			I		I			V	
MHPTS (71)	107 N-P	Medical/anesthesiology Emergency	I		A	A		A				A
NOTSS (72)	27 consultant surgeons	General/orthopedic/cardiac surgery	D									
(73)	44 P	General/orthopedic surgery	I					V		V		
NTS-NAS (13)	177 NS		D	A	I	A						
OSANTS (40)	Surgical R	General surgery	I					I		V		V
OSCAR (38)	Anesthesiologists, P-N	Emergency	I	A				V				
STAT (39)	Pediatric R and experts	Pediatric septic shock	I	D						V		A
T-NOTECHS (37)	2 surgeons, 1 intensivist, and 2 N	Trauma	I					I				A
(74)	193 HCPs	Trauma	D		A	D	V			V		
TEAM (44)	6 resuscitation experts	Emergency	D	D	D			V				V
(45)	35 HCPs	Emergency	D		A	V				V		
(46)	151 HCPs	Obstetric emergencies	I		A	V				V		
TPOT (57)	72 MS-NS	Trauma	I					V				V

The table includes the objective and population of each article. Methodological quality of the included studies by measurement properties is evaluated as very good (V), adequate (A), doubtful (D), inadequate (I) by the COSMIN method. Empty cells indicate that that measurement property (or part of it) was not performed. ANTS, Anesthetists' Non-Technical Skills; ANTSdk, Anesthesiologists' Non-Technical Skills in Denmark; ASNTS, Anesthesiology students' Non-Technical Skills; AOTP, Assessment of Obstetrical Team Performance; BARS, Behaviorally Anchored Rating Scale; CALM, Concise Assessment of Leader Management; GAOTP, Global Assessment of Obstetrical Team Performance; HCPs, Healthcare professionals; HFRS, Human Factors Rating Scale; MHPTS, Mayo High Performance Team Scale; MS, Medicine students; NOTSS, Non-Technical Skills for Surgeons; NS, Nursing students; NTS, Non-Technical skills; NTS-NAS, Non-Technical Skills-Nursing Assessment Scale; N, Nurses; O, Obstetricians; OSANTS, Objective Structured Assessment of Non-Technical Skills; OSCAR, Observational Skill-based Clinical Assessment tool for Resuscitation; Ottawa GRS, Global Rating Scale; P, Physicians; R, Residents; STAT, Simulation Team Assessment Tool; T-NOTECHS, Non-technical skills scale for trauma; TEAM, Eam, Emergency Assessment Measure; TPOT, Team Performance Observational Tool.

TABLE 3 Evaluation of the measurement properties and quality grading of the evidence of NTS instruments.

NTS instruments	Content validity		Structural validity		Internal consistency		Cross cultural validity		Reliability		Measurement error		Criterion validity		Construct validity		Responsiveness	
	Overall rating	QoE	OR	QoE	OR	QoE	OR	QoE	OR	QoE	OR	QoE	OR	QoE	OR	QoE	OR	QoE
ANTS	+	M			?	L			-	M	?	M			+	L	+	L
ANTSdk	-	VL							+	M							+	L
ASNTS	+	M							+	VL								
Assessment of EM physicians' NTS	+	H																
BARS tool	-	VL							+	M							+	L
AOTP-GAOTP	±	L			?	VL			+	L								
CALM	+	M							-	VL								
GRS	?	L			?	L	+	VL	±	M					+	L	+	L
HFRS	-	VL			?	VL			-	VL								
LOSA	-	VL			?	VL			?	VL					+	L		
MHPTS	-	VL	?	M	+	M			?	L							+	M
NOTSS	-	VL							+	L			+	L				
NTS-NAS	+	H	-	VL	?	M												
OSANTS	-	VL			?	VL			+	M					+	M	?	VL
OSCAR	+	H			?	M			+	M								
STAT	+	M							+	M					+	L		
T-NOTECHS	-	VL	+	M	?	L	?	M	±	M								
TEAM	+	M	+	H	+	H			±	M								
TPOT	-	VL			-	L			+	M					+	M		

The measurement properties of the NTS instruments are evaluated as sufficient (+), insufficient (-), inconsistent (±), or indeterminate (?) based on the coherent results obtained in one or more studies, by criteria for good measurement properties of the COSMIN method. The quality of the evidence is graded as high (H), moderate (M), low (L), very low (VL) evidence using a modified GRADE approach, based on the COSMIN method. Empty cells indicate that that measurement property (or part of it) was not performed. ANTS, Anesthetists' Non-Technical Skills; ANTSdk, Anesthesiologists' Non-Technical Skills in Denmark; AOTP, Assessment of Obstetrical Team Performance; ASNTS, Anesthesiology students' Non-Technical Skills; BARS, Behaviorally Anchored Rating Scale; CALM, Concise Assessment of Leader Management; GAOTP, Global Assessment of Obstetrical Team Performance; HFRS, Human Factors Rating Scale; MHPTS, Mayo High Performance Team Scale; NOTSS, Non-Technical Skills for Surgeons; NTS, Non-Technical skills; NTS-NAS, Non-Technical Skills-Nursing Assessment Scale; OR, Overall rating; OSANTS, Objective Structured Assessment of Non-Technical Skills; OSCAR, Observational Skill-based Clinical Assessment tool for Resuscitation; Ottawa GRS, Global Rating Scale; QoE, Quality of evidence; STAT, Simulation Team Assessment Tool; T-NOTECHS, Non-technical skills scale for trauma; TEAM, Team Emergency Assessment Measure; TPOT, Team Performance Observational Tool.

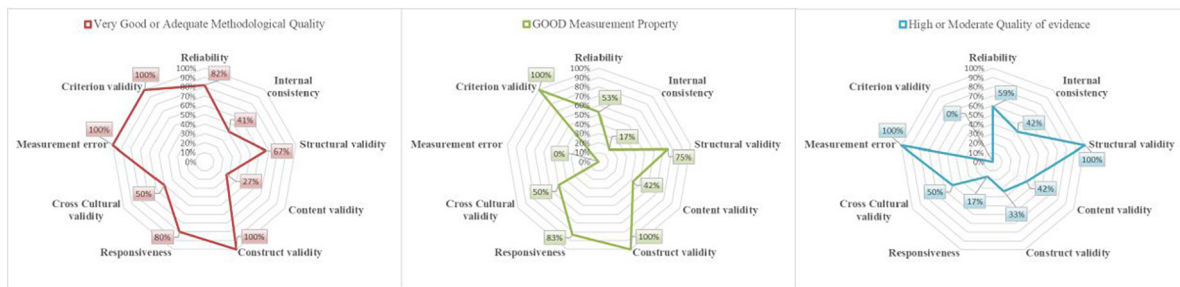


FIGURE 2

Quality assessment. The assessed measurement properties of 19 NTS instruments to which the plotted lines refer, are distributed around the radar. The red line in the first chart plots the proportion of studies with a very good or adequate methodological quality, over a total of 26 studies. The green line in the second chart represents the proportion of instruments with good measurement properties. The blue line in the third chart shows the proportion of instruments with high or moderate evidence quality. For example, content validity was of very good or adequate methodological quality in 27% of the included articles (3/11 studies), instruments with good measurement properties of content validity were 42% (8/19 instruments) and instruments with high or moderate evidence quality of content validity were 42% (8/19 instruments).

programs on the acquisition and maintenance of NTS in HCPs is essential for resource allocation and planning simulation training, including simulation content, length and frequency, and target groups. Reliable and valid NTS measurement instruments can set the stage to accurately measure change of essential CRM behaviors during simulation training.

In this study, we found that the methodological quality of instrument development was doubtful or poor for all studies, and only 1 instrument (5%) could be recommended. This finding has two main implications. First, in the domain of the assessment of NTS, instrument development and content validity studies need to rigorously report the application of a consistent methodological approach according to accredited reporting guidelines to demonstrate process validity and reliability. For this review, we used the COSMIN method as it sets clear criteria for instrument development and psychometric testing of the instruments' measuring properties using Delphi consensus based procedures. NTS assessment instruments developed before the COSMIN criteria were published (in 2010) referred to standards of prior psychometric evaluation tools. Those studies are more likely to be less compliant with the detailed and rigorous COSMIN criteria (54). Future research should aim at following recommended criteria for the evaluation of psychometric tools to safeguard the quality, validity and reliability of NTS measuring instruments for future use in the simulation setting.

Second, the research gap in this domain is wide, as there is limited evidence on essential aspects of validity and reliability for most instruments reported in this review, requiring further research. Content validity is the most important measurement property of a measuring instrument as it refers to item relevance, comprehensiveness and comprehensibility with respect to the construct of interest and study population (28). The eight

instruments (13, 17, 38, 39, 42, 44, 55, 68) that reported a positive rating of content validity should be further evaluated at least for structural validity and internal consistency to determine a recommendation for their use.

Most NTS instruments were tested for reliability. Reliability of instrument domains and items resulted lower for concepts that might be more difficult to translate into an observable behavior, such as situational awareness, teamwork or team morale (45). NTS definition through a shared framework including situated examples of expected behaviors in different settings, as reported for some instruments (17, 38, 55–57) is essential to promote users' shared understanding of the behaviors to observe and the accuracy of the application of NTS instruments.

NTS are comprehensive concepts characterized by their complexity, interconnectedness, evolving scope and meaning, irrespective of the scenario been simulated. The instruments included in this review reported primarily NTS domains described as "leadership" and "teamwork," followed by "communication" and "situation awareness." Only about half of the instruments used "decision making" and "task management," which are often included in "leadership," "teamwork" or "situation awareness." Leadership is mostly regarded in relation to managing a team or organization (58) but can also be defined as a set of personal skills or traits, or focusing on the relation between leaders and followers (59, 60). Leaders in healthcare should have both the technical and social competences to exercise effective situational leadership and a flexible approach to patient management (61). Markers of effective teamwork include: calling for help early, establishing clear roles and leadership, employing team-oriented communication techniques, establishing a team situational awareness, effective decision making, and maintaining an adequate group climate (62). On the other hand, individual and team situation awareness is a

complex dynamic process to maintain awareness of a critical clinical situation based on perception, comprehension and projection. A closer relationship with the environment, taking into account contextual factors, determines a distributed situational awareness. Finally, communication can be intended as a neutral means to share information between individuals or more as a means to structure social processes, including leadership and followership, and share mental models. Failures in communication have been found to be a leading cause of errors in healthcare determining the lack of sharing of a common mental model and understanding of the patient's conditions. Closed-loop and direct communication is an essential goal for reliability in healthcare. Establishing clear mental models on NTS is a prerequisite for content validity of instruments that evaluate NTS.

Simulation instructors, while assessing NTS, face the need to strike a balance between capturing the nuances of non-technical behavior in managing simulated patient emergencies and synthesis. ANTS, AS-NTS, AOTP/GAOTP, and TEAM were reported to have a high usability rate. Efficiency, satisfaction and effectiveness are essential domains of instrument usability (63), which should be applied to instruments devised to assess NTS. Simplicity, ease of use, and accuracy are essential to reduce assessment errors and to increase NTS evaluation during simulation training. Moreover, the NTS instruments included in this review were developed and used for emergency scenarios in different clinical settings. Some instruments present examples of expected NTS behavior or are directed to specific teams, to increase usability and transferability in specific settings (38, 40, 43, 64).

Limitations

This review has some limitations. We included instruments for the evaluation of NTS during simulation training with the exclusion of instruments devised for observing NTS in the clinical setting. While the simulation setting is where HCPs and healthcare students' NTS training takes place, the clinical setting is where NTS ultimately should be practiced. In order to evaluate the effect of CRM simulation training on healthcare practices, NTS instruments should be validated both on simulation and real settings to be able to compare and evaluate the uptake of those skills.

Conclusion

Out of a large variety of published instruments devised to assess NTS, the TEAM instrument can be recommended for use during HFS. Evidence is still limited on essential aspects of

validity and reliability of all the other NTS instruments included in this review. Further research is needed to establish their performance in order to be reliably used for HFS.

Data availability statement

The original contributions presented in this study are included in the article/**Supplementary material**, further inquiries can be directed to the corresponding author.

Author contributions

OG conceived and designed the study. OG, KRT, CG, SE, VS, IE, and FC carried out the systematic literature review. OG, KRT, CG, SE, VS, AV, and FC involved in data analysis and performed the quality assessment. OG, KRT, CG, VS, and FC drafted the manuscript. All authors revised the manuscript and approved the final version as submitted.

Funding

This work received funding by the Italian Ministry of Health with 5x1000 funds 2022ET.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmed.2022.986296/full#supplementary-material>

References

- Makary MA, Daniel M. Medical error—the third leading cause of death in the US. *BMJ*. (2016). 353:i2139. doi: 10.1136/bmj.i2139
- Institute of Medicine (US) Committee on Quality of Health Care in America. In: Kohn LT, Corrigan JM, Donaldson MS, editors. *To Err is Human: Building a Safer Health System*. Washington (DC): National Academies Press (2000).
- World Health Organization [WHO]. *Patient Safety: Making Health Care Safer*. Geneva: WHO (2017).
- Gawande AA, Zinner MJ, Studdert DM, Brennan TA. Analysis of errors reported by surgeons at three teaching hospitals. *Surgery*. (2003) 133:614–21.
- Müller MP, Hänsel M, Stehr SN, Fichtner A, Weber S, Hardt F, et al. Six steps from head to hand: a simulator based transfer oriented psychological training to improve patient safety. *Resuscitation*. (2007) 73:137–43. doi: 10.1016/j.resuscitation.2006.08.011
- St-Pierre M, Hofinger G. *Human Factors und Patientensicherheit in der Akutmedizin*. 3rd ed. Berlin: Springer (2014). p. 378.
- Hoffmann B, Siebert H, Euteneier A. Patient safety in education and training of healthcare professionals in Germany. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz*. (2015) 58:87–94.
- Nassief K, Azer M, Watts M, Tuala E, McLennan P, Curtis K. Emergency department care-related causal factors of in-patient deterioration. *Aust Health Rev*. (2021) 46:35–41. doi: 10.1071/AH21190
- Aaberg OR, Hall-Lord ML, Husebø SIE, Ballangrud R. A human factors intervention in a hospital - evaluating the outcome of a TeamSTEPPS program in a surgical ward. *BMC Health Serv Res*. (2021) 21:114. doi: 10.1186/s12913-021-06071-6
- Moreno RP, Rhodes A, Donchin Y. Patient safety in intensive care medicine: the declaration of Vienna. *Intensive Care Med*. (2009) 35:1667–72.
- Flin R, Maran N. Basic concepts for crew resource management and non-technical skills. *Best Pract Res Clin Anaesthesiol*. (2015) 29:27–39. doi: 10.1016/j.bpa.2015.02.002
- Kim J, Neilipovitz D, Cardinal P, Chiu M, Clinch J. A pilot study using high-fidelity simulation to formally evaluate performance in the resuscitation of critically ill patients: the university of ottawa critical care medicine, high-fidelity simulation, and crisis resource management. *Crit Care Med*. (2006) 34:2167–74. doi: 10.1097/01.CCM.0000229877.45125.CC
- Pires SMP, Monteiro SOM, Pereira AMS, Stocker JNM, Chaló DM, Melo EMOP. Non-technical skills assessment scale in nursing: construction, development and validation. *Rev Lat Am Enfermagem*. (2018) 26:e3042. doi: 10.1590/1518-8345.2383.3042
- Gross B, Rusin L, Kiesewetter J, Zottmann JM, Fischer MR, Prückner S, et al. Crew resource management training in healthcare: a systematic review of intervention design, training conditions and evaluation. *BMJ Open*. (2019) 9:e025247. doi: 10.1136/bmjopen-2018-025247
- Mellin-Olsen J, Staender S, Whitaker DK, Smith AF. The helsinki declaration on patient safety in anaesthesiology. *Eur J Anaesthesiol*. (2010) 27:592–7.
- Flin R, O'Connor P, Crichton M. *Safety at the Sharp End: A Guide to Non-Technical Skills*. London: CRC Press (2017). p. 330.
- Moll-Khosrawi P, Kamphausen A, Hampe W, Schulte-Uentrop L, Zimmermann S, Kubitz JC. Anaesthesiology students' Non-Technical skills: development and evaluation of a behavioural marker system for students (AS-NTS). *BMC Med Educ*. (2019) 19:205. doi: 10.1186/s12909-019-1609-8
- Gaba DM, Fish KJ, Howard SK. *Crisis Management in Anesthesiology*. New York, NY: Churchill Livingstone (1994). p. 309.
- Clay-Williams R, Braithwaite J. Determination of health-care teamwork training competencies: a Delphi Study. *Int J Qual Health Care*. (2009) 21:433–40.
- McIlhenny C, Yule S. Learning non-technical skills through simulation. In: Biyani CS, Van Cleynenbreugel B, Motttrie A, editor. *Practical Simulation in Urology*. Cham: Springer (2022). p. 289–305.
- Gaba DM. The future vision of simulation in health care. *Qual Saf Health Care*. (2004) 13(Suppl_1):i2–10.
- Cheng A, Magid DJ, Auerbach M, Bhanji F, Bigham BL, Blewer AL, et al. Part 6: resuscitation education science: 2020 American heart association guidelines for cardiopulmonary resuscitation and emergency cardiovascular care. *Circulation*. (2020) 142(16_Suppl_2):S551–79. doi: 10.1161/CIR.0000000000000903
- Wyckoff MH, Singletary EM, Soar J, Olasveengen TM, Greif R, Liley HG, et al. International consensus on cardiopulmonary resuscitation and emergency cardiovascular care science with treatment recommendations. *Resuscitation*. (2021) 169:229–311.
- INACSL Standards Committee. INACSL standards of best practice: SimulationSM?: operations. *Clin Simul Nurs*. (2017) 13:681–7.
- Downing D, Chang TP, Robertson JM, Anderson M, Diaz DA, Spain AE. Terminology and concepts working group. Second ed. In: Lioce L, editor. *Healthcare Simulation Dictionary*. Rockville, MD: Agency for Healthcare Research and Quality (2020).
- Patterson MD, Geis GL, LeMaster T, Wears RL. Impact of multidisciplinary simulation-based training on patient safety in a paediatric emergency department. *BMJ Qual Saf*. (2013) 22:383–93. doi: 10.1136/bmjqs-2012-000951
- Figueroa MI, Sepanski R, Goldberg SP, Shah S. Improving teamwork, confidence, and collaboration among members of a pediatric cardiovascular intensive care unit multidisciplinary team using simulation-based team training. *Pediatr Cardiol*. (2013) 34:612–9. doi: 10.1007/s00246-012-0506-2
- Bank I, Snell L, Bhanji F. Pediatric crisis resource management training improves emergency medicine trainees' perceived ability to manage emergencies and ability to identify teamwork errors. *Pediatr Emerg Care*. (2014) 30:879–83. doi: 10.1097/PEC.0000000000000302
- Garvey AA, Dempsey EM. Simulation in neonatal resuscitation. *Front Pediatr*. (2020) 8:59. doi: 10.3389/fped.2020.00059
- Boet S, Larrigan S, Martin L, Liu H, Sullivan KJ, Etherington C. Measuring non-technical skills of anaesthesiologists in the operating room: a systematic review of assessment tools and their measurement properties. *Br J Anaesth*. (2018) 121:1218–26. doi: 10.1016/j.bja.2018.07.028
- Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. (2021) 372:n71.
- Palaganas JC, Maxworthy JC, Epps CA, Mancini ME. *Defining Excellence in Simulation Programs*. China: Wolters Kluwer (2014).
- Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. *Syst Rev*. (2016) 5:210. doi: 10.1186/s13643-016-0384-4
- Mokkink LB, de Vet HCW, Prinsen CAC, Patrick DL, Alonso J, Bouter LM, et al. COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. *Qual Life Res*. (2018) 27:1171–9.
- Prinsen CAC, Mokkink LB, Bouter LM, Alonso J, Patrick DL, de Vet HCW, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res*. (2018) 27:1147–57.
- Terwee CB, Prinsen CAC, Chiarotto A, Westerman MJ, Patrick DL, Alonso J, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi Study. *Qual Life Res*. (2018) 27:1159–70.
- Steinemann S, Berg B, DiTullio A, Skinner A, Terada K, Anzelon K, et al. Assessing teamwork in the trauma bay: introduction of a modified "NOTECHS" scale for trauma. *Am J Surg*. (2012) 203:69–75. doi: 10.1016/j.amjsurg.2011.08.004
- Walker S, Brett S, McKay A, Lambden S, Vincent C, Sevdalis N. Observational skill-based clinical assessment tool for resuscitation (OSCAR): development and validation. *Resuscitation*. (2011) 82:835–44. doi: 10.1016/j.resuscitation.2011.03.009
- Reid J, Stone K, Brown J, Caglar D, Kobayashi A, Lewis-Newby M, et al. The simulation team assessment tool (STAT): development, reliability and validation. *Resuscitation*. (2012) 83:879–86. doi: 10.1016/j.resuscitation.2011.12.012
- Dedy NJ, Szasz P, Louridas M, Bonrath EM, Husslein H, Grantcharov TP. Objective structured assessment of nontechnical skills: reliability of a global rating scale for the in-training assessment in the operating room. *Surgery*. (2015) 157:1002–13. doi: 10.1016/j.surg.2014.12.023
- Watkins SC, Roberts DA, Boulet JR, McEvoy MD, Weinger MB. Evaluation of a simpler tool to assess nontechnical skills during simulated critical events. *Sim Healthcare*. (2017) 12:69–75. doi: 10.1097/SIH.0000000000000199
- Fletcher G, Flin R, McGeorge P, Glavin R, Maran N, Patey R. Anaesthetists' non-technical skills (ANTS): evaluation of a behavioural marker system. *Br J Anaesth*. (2003) 90:580–8.
- Tregunno D, Pittini R, Haley M, Morgan PJ. Development and usability of a behavioural marking system for performance assessment of obstetrical teams. *Qual Saf Health Care*. (2009) 18:393–6. doi: 10.1136/qshc.2007.026146
- Cooper S, Cant R, Porter J, Sellick K, Somers G, Kinsman L, et al. Rating medical emergency teamwork performance: development of the team emergency assessment measure (TEAM). *Resuscitation*. (2010) 81:446–52. doi: 10.1016/j.resuscitation.2009.11.027

45. Freytag J, Stroben F, Hautz WE, Schaub SK, Kämmer JE. Rating the quality of teamwork—a comparison of novice and expert ratings using the team emergency assessment measure (TEAM) in simulated emergencies. *Scand J Trauma Resusc Emerg Med.* (2019) 27:12. doi: 10.1186/s13049-019-0591-9
46. Carpini JA, Calvert K, Carter S, Epee-Bekima M, Leung Y. Validating the team emergency assessment measure (TEAM) in obstetric and gynaecologic resuscitation teams. *Aust N Z J Obstet Gynaecol.* (2021) 61:855–61.
47. Ginsburg LR, Tregunno D, Norton PG, Smee S, de Vries I, Sebok SS, et al. Development and testing of an objective structured clinical exam (OSCE) to assess socio-cultural dimensions of patient safety competency. *BMJ Qual Saf.* (2015) 24:188–94.
48. Levett-Jones T, Dwyer T, Reid-Searl K, Heaton L, Flenady T, Applegarth J, et al. *Patient Safety Competency Framework (PSCF) for Nursing Students.* Queensland: USC Research Bank (2017).
49. Heinen M, van Oostveen C, Peters J, Vermeulen H, Huis A. An integrative review of leadership competencies and attributes in advanced nursing practice. *J Adv Nurs.* (2019) 75:2378–92.
50. Okuyama A, Martowirono K, Bijnen B. Assessing the patient safety competencies of healthcare professionals: a systematic review. *BMJ Qual Saf.* (2011) 20:991–1000.
51. Kirkman MA, Sevdalis N, Arora S, Baker P, Vincent C, Ahmed M. The outcomes of recent patient safety education interventions for trainee physicians and medical students: a systematic review. *BMJ Open.* (2015) 5:e007705.
52. Alken A, Luursema JM, Weenk M, Yauw S, Fluit C, van Goor H. Integrating technical and non-technical skills coaching in an acute trauma surgery team training: is it too much? *Am J Surg.* (2018) 216:369–74.
53. Urbina J, Monks SM. *Validating Assessment Tools in Simulation.* Treasure Island, FL: StatPearls Publishing (2022).
54. Rosenkoetter U, Tate RL. Assessing features of psychometric assessment instruments: a comparison of the COSMIN checklist with other critical appraisal tools. *Brain Impairment.* (2018) 19:103–18.
55. Flowerdew L, Brown R, Vincent C, Woloshynowych M. Development and validation of a tool to assess emergency physicians' nontechnical skills. *Ann Emerg Med.* (2012) 59:376–85.e4.
56. Hull L, Arora S, Kassab E, Kneebone R, Sevdalis N. Observational teamwork assessment for surgery: content validation and tool refinement. *J Am Coll Surg.* (2011) 212: 234-43.e1–5.
57. Zhang C, Miller C, Volkman K, Meza J, Jones K. Evaluation of the team performance observation tool with targeted behavioral markers in simulation-based interprofessional education. *J Interprof Care.* (2015) 29:202–8.
58. Gosling J, Mintzberg H. *The Five Minds of a Manager.* Brighton, MA: Harvard Business Review (2003).
59. Metcalfe BA, Alban J. Leadership in public sector organizations. In: Storey J editor. *Leadership in Organizations.* Milton Park: Routledge (2003).
60. Bolden R. *What is Leadership?* Exeter: University of Exeter Centre of Leadership Studies. (2020).
61. Steinert Y, Naismith L, Mann K. Faculty development initiatives designed to promote leadership in medical education. A BEME systematic review: BEME Guide No. 19. *Med Teach.* (2012) 34:483–503.
62. Brady PW, Giambra BK, Sherman SN, Clohessy C, Loechtenfeldt AM, Walsh KE, et al. The parent role in advocating for a deteriorating child: a Qualitative Study. *Hosp Pediatr.* (2020) 10:728–42.
63. Healthcare Information Management Systems Society. *Nine Essential Principles of Software Usability for EMRs.* Chicago, IL: HIMSS (2015).
64. Morgan PJ, Pittini R, Regehr G, Marrs C, Haley MF. Evaluating teamwork in a simulated obstetric environment. *Anesthesiology.* (2007) 106:907–15.
65. Jirativanont T, Raksamani K, Aroonpruksakul N, Apidechakul P, Suraseranivongse S. Validity evidence of non-technical skills assessment instruments in simulated anaesthesia crisis management. *Anaesth Intensive Care.* (2017) 45:469–75.
66. Jepsen RMHG, Spanager L, Lyk-Jensen HT, Dieckmann P, Østergaard D. Customisation of an instrument to assess anaesthesiologists' non-technical skills. *Int J Med Educ.* (2015) 6:17–25.
67. Jepsen RMHG, Dieckmann P, Spanager L, Lyk-Jensen HT, Konge L, Ringsted C, et al. Evaluating structured assessment of anaesthesiologists' non-technical skills. *Acta Anaesthesiol Scand.* (2016) 60:756–66.
68. Nadkarni LD, Roskind CG, Auerbach MA, Calhoun AW, Adler MD, Kessler DO. The development and validation of a concise instrument for formative assessment of team leader performance during simulated pediatric resuscitations. *Sim Healthcare.* (2018) 13:77–82.
69. Franc JM, Verde M, Gallardo AR, Careno L, Ingrassia PL. An Italian version of the Ottawa crisis resource management global rating scale: a reliable and valid tool for assessment of simulation performance. *Intern Emerg Med.* (2017) 12:651–6.
70. Moorthy K, Munz Y, Adams S, Pandey V, Darzi A. A human factors analysis of technical and team skills among surgical trainees during procedural simulations in a simulated operating theatre. *Ann Surg.* (2005) 242:631–9.
71. Malec JF, Torsher LC, Dunn WF, Wiegmann DA, Arnold JJ, Brown DA, et al. The mayo high performance teamwork scale: reliability and validity for evaluating key crew resource management skills. *Simul Healthc.* (2007) 2:4–10.
72. Yule S, Flin R, Paterson-Brown S, Maran N, Rowley D. Development of a rating system for surgeons' non-technical skills. *Med Educ.* (2006) 40:1098–104.
73. Yule S, Flin R, Maran N, Rowley D, Youngson G, Paterson-Brown S. Surgeons' non-technical skills in the operating room: reliability testing of the NOTSS behavior rating system. *World J Surg.* (2008) 32:548–56.
74. Repo JP, Rosqvist E, Lauritsalo S, Paloneva J. Translatability and validation of non-technical skills scale for trauma (T-NOTECHS) for assessing simulated multi-professional trauma team resuscitations. *BMC Med Educ.* (2019) 19:40. doi: 10.1186/s12909-019-1474-5