SCOTTISH MEDICAL JOURNAL

Original Article

Scottish Medical Journal
0(0) 1–10

DOI: 10.1177/00369330211008594
journals.sagepub.com/home/scm
SAGE

Capturing the non-technical skills of a technical skills trainer (NTS-TeST) during simulation

Matthew Pears[1,*], Chandra Shekhar Biyani[2,*], Adrian D Joyce[3], Ken Spearpoint[4], Marina Yiasemidou[5], Ben Van Cleynenbreugel[6], Jake Patterson[7] and Faisal Mushtaq[8]

Abstract

Objective: To develop an assessment instrument that can be used as a comprehensive feedback record to convey to a trainer the non-technical aspects of skill acquisition and training.

Methods: The instrument was developed across three rounds. In Round 1, 6 endourological consultants undertook a modified Delphi process. Round 2 included 10 trainers who assessed each question's relevance and practicability. Round 3 involved a pilot study with fifteen urology residents who participated in a technical skills simulation session with the incorporation of the instrument. We report the content, face, and construct validity, and the internal consistency of an NTS instrument for trainers.

Results: The instrument had a consistent and a high positive average for each of the 4 sections of the instrument, regardless of the type of user. Positive Spearman's correlation coefficients (0.02 to .64) for content validity and Cronbach's alpha ($a = 0.70$) indicated good validity and moderate reliability of the instrument.

Conclusion: We propose a novel NTS instrument for trainers during a simulation. This instrument can be used for benchmarking the quality of technical skills simulation training.

Keywords

Simulation, training, trainers, nontechnical skills, trainees, technical skills

Introduction

Technical skills simulation has become a pre-requisite for educational course completion in healthcare.[1] To achieve a high efficacy of technical skills acquisition from simulation, a trainer is positioned to optimise each simulation session with trainees.

Teamwork and communication are vital during simulation training to pass on information. Therefore, the pedagogical characteristics of a trainer must include excellent abilities to verbalise their tacit knowledge, and have compensatory behaviours that adapt to the needs and shortcomings of trainees in real time.[2] Yet, there is no standardised and widely used instrument to capture this phenomenon to ensure this process is occurring. The evaluation of a trainee's feedback of their trainer's technical skills teaching competencies is vital. Similar instruments that do exist tend to have poor reliability and validity testing.[3]

There is a need for such an instrument to allow students to evaluate the performance of their trainers, as although experts may have excellent levels of performance, status, and creativity in practice when complex actions are performed, there is compression of cognitive and behavioural processes leading to

[1]Applied Cognition and Healthcare Researcher, School of Psychology, University of Leeds, UK
[2]Consultant Urologist, St James University Hospital, UK
[3]Honorary Consultant Urologist, St James University Hospital, UK
[4]Principal Lecturer, Programme Lead, MSc Health & Medical Simulation University of Hertfordshire, Hatfield, UK
[5]NIHR Academic Clinical Lecturer in General Surgery, Academic Surgery, Hull York Medical School, ST6 Colorectal Surgery, Bradford Teaching Hospitals, UK
[6]Urologist, University Hospitals, Leuven, Belgium
[7]Consultant Urologist, Royal Hallamshire Hospital, Sheffield, UK
[8]Associate Professor, School of Psychology, University of Leeds, UK

[*]These authors contributed equally to this article.

Corresponding author:
Chandra Shekhar Biyani, Department of Urology, St James's University Hospital, Beckett Street, Leeds LS9 7TF, UK.
Email: shekharbiyani@hotmail.com

automated and shortcut actions that can be hard to decompress when describing, or training others.[4] Due to this difficulty in teaching, attempts to implement mechanisms that capture this information during a technical skills simulation have had mixed success.[5] One informative gastrointestinal endoscopy training review provided an initial insight into a trainer's effects on learning experiences of students.[6] They found that with a training framework in place that incorporated performance enhancing feedback, trainers can be better verbalise deconstructed technical skills. Also the introduction of 'Training the Trainer' programmes to specialty training curricula aimed to improve general teaching skills, including communication, and incorporated such models as Kolb's learning cycle.[7] However, these attempts either used self-reflection and subjective feedback measures, or were not specifically focused on the assessment of appropriate learning outcomes relating to the communication and teamworking skills of the trainer in this context. We have developed such an instrument to address this absence of a non-technical skills (NTS) instrument that allows trainees to assess a trainer during technical skills simulation.

Method

A 3-round project was designed to develop and validate an assessment instrument for non-technical skills of trainers during technical skills simulation (NTS-TeST). Using a modified Delphi process, round 1 formed expert feedback and formation of the assessment instrument. Round 2 measured the content validity. Round 3 was a pilot test of the instrument and evaluated its internal consistency and construct validity (Figure 1).

Round 1 – Delphi process and derivation of the instrument

Several surgical training behaviours were considered based upon a technical skills measurement instrument formerly created.[8] This previous instrument ensured trainee satisfaction of their trainers' NTS teaching and further modifications were finalised by the experts to be specific for use in simulation training. The published tool was designed for live surgical training and consisted of 27 questions that may not be relevant/necessary during technical skills simulation session, in addition, modification was necessary to improve wording/structure to fit with simulation settings. Content validity (the extent to which the questions in an instrument represent the theoretical construct under investigation) was asserted by concurrence from 6 subject matter experts ensuring the instrument covered the main NTS components of interest. Subject matter experts were recruited during the urology simulation boot camp course.[9] They were identified by sociometric analysis (who do others ask for help), extended domain experience (deliberate practice time), performance analysis (high positive record), and role within their institution. These individuals were viewed as being able to effectively communicate their knowledge to trainees. Each expert had more than 5 years of consultant experience and pedagogical responsibilities. Communication was through online discussion and email correspondence over a 1-month period asked them to openly suggest the elimination or addition of a question to the instrument list and to make comments for why this was their opinion. Final responses about the relevance of each sub-category were captured on a 5-point scale (1="very relevant and succinct", 2="relevant but needs minor alterations", 3="unable to assess relevance", 4="not relevant unless major alterations", 5="not relevant"). The final instrument had 4 sections that each represented a different time point (before, during, after and global rating scale) in a trainee's simulation session that experts had agreed captured key underpinnings of a trainer's NTS.

Round 2 – Trainers' review

To further evaluate the instrument, 10 trainers with at least 5 years' experience in teaching technical skills were asked for input on the accuracy in representing their NTS. This round allowed us to understand if the instrument would be accepted by trainers who would be incorporating it when teaching their surgical students. Their responses about the relevance of each sub-category were captured on the same 5-point scale used by the round 1 experts. Additionally, feedback on item descriptions, content, format, wording of items (e.g., double-barrelled, or leading questions) and item similarities were collected. Reiteration of these changes was achieved by both quantitative ratings and open-ended comments, which allowed flexible communication towards modification of the final 18 question NTS-TeST version (Tables 1 and 2).

Round 3 – Piloting with trainees

Round 3 involved piloting the instrument with 15 surgical trainees during supervised technical skills simulation. It is important to note they were not rating their trainer but rating the questions with the same 5-point Likert scale as in round 1 and 2, by acknowledging if each question reflected their trainer's abilities. This was after they had completed a 4-hour surgical technical skills simulation session. Face validity was investigated as the surgical trainees were asked for feedback on the NTS-TeST to ensure they agreed they could use it to measure their trainer's NTS during training. The goal

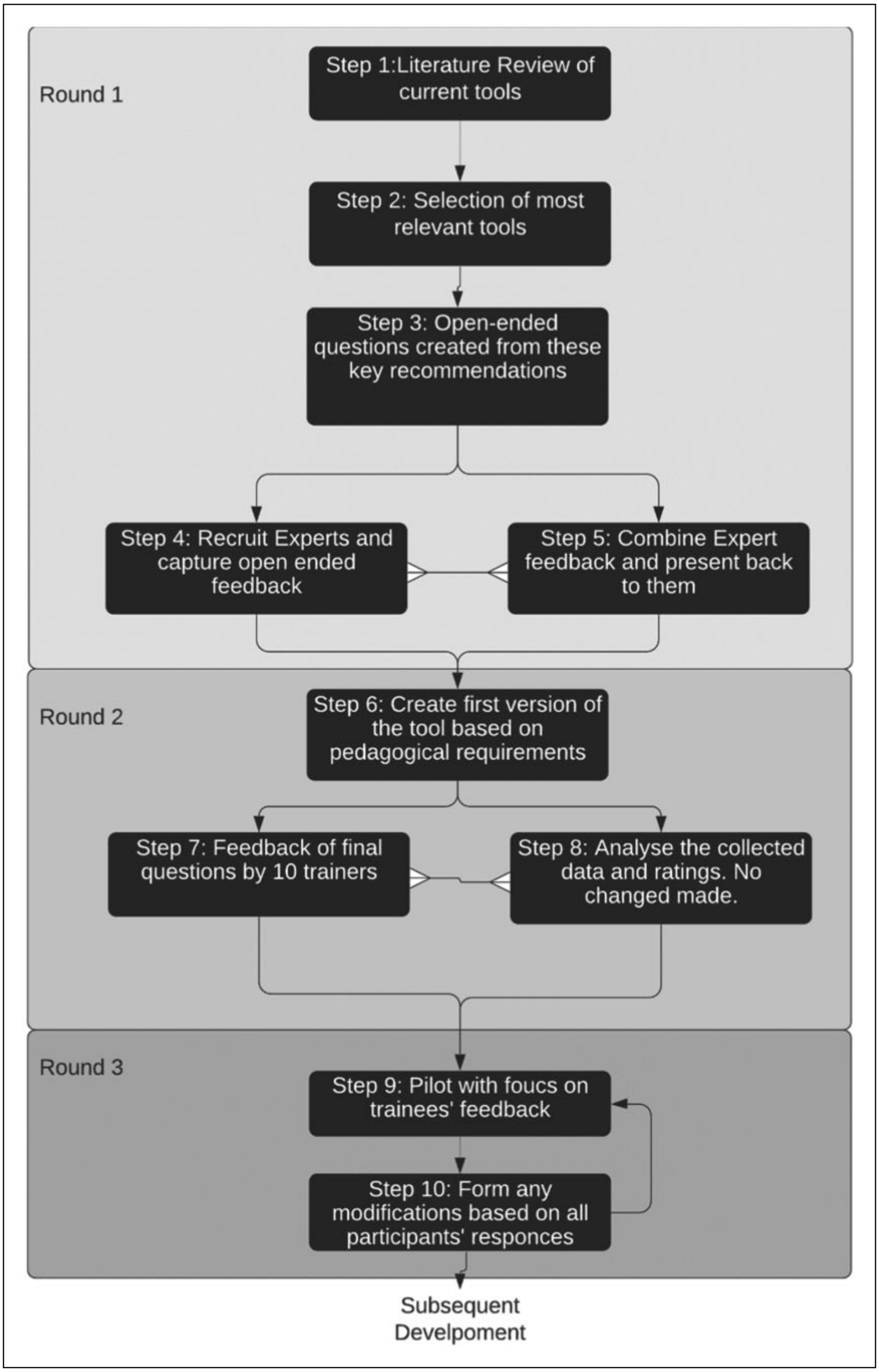

Figure 1. The 3 rounds performed in the modified Delphi process, broken down into the 10 methodological steps performed by the researchers.

was to assess if the 4 sections of the instrument were compiled correctly and to explore each section for any confounding factors which would impact on true responses from the intended group of raters, being the surgical trainees.

Analysis

Construct validity was captured by performing a Spearman's correlation coefficient test. The threshold value for the correlation coefficient was set at .5 as this is generally accepted in validation studies.[10,11] The

Table 1. The final modified scale assessment of non-technical skills of a trainer during simulated technical skills training.

Stage	Skill	1	2	3	4	5
Pre-skill training	Q1 Preparation	None		Adequately checked necessary kit		Well prepared
	Q2 Trainees expectations	Nil check		Adequately explored		Well established
	Q3 Explanation about the session	No explanation (if applicable)		Relevant aspects of surgery explained adequately		Relevant aspects of surgery explained clearly
	Q4 Aims	None		Aims of trainee outlined adequately		Aims of trainee outlined clearly
	Q5 Trainer support during training	No guidelines (if applicable)		Reasonable guidelines when trainer will take over from trainee to demonstrate steps		Precise guidelines when trainer will take over from trainee to demonstrate steps
During skill training	Q6 Instructions	Poor, not consistent, use vague terminology during teaching		Reasonably clear with some consistent instructions. Adequate use of common language. Some deconstruction of steps		Clear and consistent instructions. Standardised language. Deconstruct steps
	Q7 Assistance	Poor assistance, poor concentration during session		Adequate assistance		Good and helpful assistance
	Q8 Time	Pressure on trainee on time constraint and to hurry procedure		Reasonable pressure on trainee on time constraint/hurry procedure		No pressure on trainee on time constraint or to hurry procedure
	Q9 Errors	Does not allow trainee to continue to operate after 1 minor technical error		Allows trainee to continue to operate after a few minor technical errors		Allows trainee to continue to operate after several minor technical errors or 1 major error.
	Q10 Operative field	Trainer constantly interfering into surgical field		Trainer only occasionally interferes into surgical field		Trainer does not interfere actively into surgical field
	Q11 Supervision	Dictatorial supervision of trainee and negative comments		Allows some autonomy of trainee without comprising patient safety, some positive and negative comments		Allows complete autonomy of trainee without comprising patient safety and positive encouragement
	Q12 Explanation of technical tasks	No explanation (if applicable)		Explains adequately technical tasks during procedure to trainee (if applicable)		Explains clearly technical tasks during procedure to trainee (if applicable)
	Q13 Demonstration of technical tasks	No helpful demonstration (if applicable)		Demonstrates technical tasks adequately without significant interfering with surgery (if applicable)		Demonstrates technical tasks clearly without significant interfering with surgery (if applicable)
Post-skill training	Q14 Feedback	No feedback		Adequate feedback on surgical strengths and/or areas of improvement to trainee		Clear feedback on surgical strengths and/or areas of improvement to trainee
	Q15 Recommendations	No future outline		Adequately outlines future surgical goals and objectives to trainee		Clearly outlines future surgical goals and objectives to trainee
Global Rating Scale trainer's behaviour	Q16 Allows trainee to ask questions	Never		Occasionally		Always
	Q17 Interrogating	Never		Occasionally		Always
	Q18 Encouraging	Never		Occasionally		Always

Table 2. Table showing underlying non-technical skills for each question inclusion, aiding rationalisation to finalise these questions into the final question list.

Question content	Trainee skill being modified	Non-technical skill of interest
Q1 Preparation including simulator	Task Familiarization	Stronger prepared mental model/schema
Q2 Trainees expectations	Event Familiarization	Stronger prepared mental model/schema
Q3 Explanation about the session	Process Familiarization	Stronger prepared mental model/schema
Q4 Aims	Task planning to the end goal	Stronger prepared mental model/schema
Q5 Support	When to watch a task	Level 1/2 Situational Awareness- perception/comprehension
Q6 Instructions	Task analysis to reach goal	Level 1/2 Situational Awareness- perception/comprehension
Q7 Assistance	Avoid inability to process information, due to excess	Cognitive Load/Overload management
Q8 Time	Allows more information gathering and decision-making time	Level 3 Situational Awareness- prediction
Q9 Errors	Consequence of Action	Decision-making, awareness/effects of actions
Q10 Operative field	Information for decision-making	Level 2 Situational Awareness- perception
Q11 Supervision	Allows exploration with safety of intervention if action not correct	Cognitive Overload and error prevention
Q12 Explanation of technical tasks	Task analysis	Level 1/2 Situational Awareness- perception/comprehension
Q13 Demonstration of technical skills	Reduce cognitive load by pre-loading schema/scripts of tasks	Cognitive Load/Overload management
Q14 Feedback	Meta- Reflection of performance and decisions	Stronger future mental model/schema
Q15 Recommendations	General trainer feedback	General trainer feedback
Q16 Allows to ask	General trainer feedback	Information gathering
Q17 Interrogating	General trainer feedback	Awareness of events
Q18 Encouraging	General trainer feedback	Motivation to continue desired behaviours

minimum sample size required for reliable Spearman's correlation coefficient was 29.[12] Internal consistency was also assessed to determine the reliability of each question towards capturing trainers' NTS in training. Cronbach's alpha was performed with the assumption that all questions attempt to measure NTS based on the previous content and construct validity fulfilled in rounds 1 and 2.[13,14] Because this study investigated the first NTS-TeST version (the first round of effects of modifications), a mix in correlative strength between questions was expected. This meant that when some questions increase in scores, other questions may not increase score at the same rate if they had low correlation. These processes enabled thorough analysis of reliability of all questions towards capture of a theoretical component (Cronbach's alpha), but also analysis of validity to ensure they are measuring the intended factor of trainers' NTS (Spearman's coefficient).

Results

The Mode for Q9 (Error) was 4, all other questions had a mode of 5. All questions had a median of 5 This meant data were skewed with range from −1 to −3.3 which may have been because of sample size. Larger standard deviations and range of responses for Q15 (Recommendations) and Q17 (Interrogating) may be initial sign of undesired variance in agreement of what these questions aim to measure.

Validity

Cohen's d was used to evaluate the strength of the relationships, where coefficients between .1 and .29 represent a small effect size, coefficients between .30 and .49 represent a moderate effect size, and coefficients above .50 indicate a large effect size.[15] The result of the 18x18 correlation matrix suggested that most questions positively correlate with each other. Questions 10 (Operative Field), 12 (Explanation of technical tasks), and 15 (Recommendations) had weak negative correlations with several other questions, which were explored with Cronbach's alpha and in the discussion. When cut-off for Spearman's coefficient was =>.1 (indicating at least a weak correlative strength), approximately 80% of total positive correlations were accounted for. When cut-off was =>.3 (indicating at least a moderate correlative strength), approximately 30% of total positive correlations were found to be in this category. Several correlations had a large =>.5 result. For example, for Q8 (Time) and Q9 (Error) (p<.001, coefficient=.626), there was a strong positive correlation. This concurred with the expert's underpinning assumptions of the constructs of these 2 questions. This also permitted the statement that within the instrument, most questions were either weak, moderate, or strongly correlated with each other towards measuring the overall factor of NTS, as concurred by those in round 1 and 2 of the study.

Reliability

The Cronbach's alpha coefficient was evaluated using the guidelines suggested by George and Mallery[16] where >.9 excellent, >.8 good, >.7 acceptable, >.6 questionable, >.5 poor, and $\leq$ .5 unacceptable. To find the overall alpha, all 18 items for NTS-TeST gained a Cronbach's alpha coefficient of 0.73, indicating acceptable reliability. Q14 (Feedback) was negatively correlated with NTS. This variable was reverse coded to improve reliability.

To explore where the NTS-TeST could be strengthened to increase the alpha level and correlations between each question, questions with similar underpinning NTS were analysed as a group. Questions relating to preparation were analysed and this group had an alpha coefficient of .729. Questions relating to situational awareness were analysed as a group. This group had an alpha coefficient of .599. Questions 7, 11, and 13 were included as capture of Cognitive Load management by the trainer and were analysed as a group. This group had an alpha coefficient of .596. With their respective Cronbach's alpha being under .7, there may be reliability issue with one or more questions within 6–13 (Table 3).

To pinpoint the issue causing moderate internal consistence, 'Cronbach's alpha if deleted' (ad) was used to identify which question was weakest in reliability. This removes a potentially weaker question from its formula and ascertains the new alpha without that question. In addition, 'corrected item-total correlation' (citc) was used and is the outcome of correlating a question with the sum of the remaining questions. It has similarity to Spearman's correlation coefficient, as it also showed which high scores on the NTS-TeST correlate positively and strongly with another high score for each specific question.

Questions 15 (recommendations) had several weak negative correlations and have low correlation with the rest of the questions (citc=.088). If removed from the instrument it would produce marginal improvement to the instrument's reliability (ad= .702). This increase in Cronbach's alpha was very small, and although the question may be more technically orientated, as it refers to recommended technical changes and removal of the question may cause a gap in the understanding of trainers' ability to provide clear summative feedback.

Question 17 (interrogation) showed the highest alpha after deletion (a=.730) compared to all other questions. However, when compared to the other questions, Q17 showed it had low interrelatedness but was not the lowest (citc=.158) and had several moderate positive Spearman's coefficient correlation results. This meant that this question did reduce the instruments reliability and validity, but not substantially. It captures some relevant information that pertains to the goal of the instrument therefore may be kept and revision to wording was suggested (see below).

Face validity and usage

We also undertook qualitative analysis to explore whether the visuals and design could be enhanced. As the Likert scale was in a portrait layout options 2 and 4 had no description. Although participants understood all 5 options, their feedback suggested the NTS-TeST to be in a landscape layout for improved equal presentation of the Likert scale options. An optional 'Name:' and 'Email:' was added if participants would like a copy of their completed instrument. Lastly, for the global measure of Q17 (Interrogation) the mean was comparatively low (3) in addition to a large range between 1 and 5. This highlighted that some participants were unable to assess relevance of this item, and it was noticed that for some participants this item was either viewed very relevant and succinct, or not relevant. Trainees' comments on their varied feedback on 'Interrogation', identified that using this term was viewed negatively by some as it was felt to denote negatively critical examination, where other trainees understood the element it intended to capture was not negative but constructive in its critical nature.

Discussion

The aim of this study was to develop and pilot an instrument that could be used by trainees to measure and assess their trainers' NTS during an educational technical skills simulation. The overarching objective was to provide surgical trainers with objective feedback to safely inform them and promote critical reflection. This was on the premise that this should help them improve their own performance in delivering and encouraging NTS.

After modifications to a previous instrument, the development, testing, and efficacy of a novel NTS-TeST suggested it had good validity and moderate internal consistency towards measuring trainers' usage of skills, as experienced by trainees during technical skills simulation. The results also showed that expert insight into such a development enabled consistent agreement on the underpinning of NTS being

Table 3. Categories of implementation objectives and questions related to categories (adapted and modified from (63, 64)).

Objective	Comment
Reaction and planned action	This defines the level of reaction and satisfaction you want to achieve with both the target audience and major stakeholders. To consider the following questions • How important is user and stakeholder acceptance to you? • Aside from the simulated procedure, how difficult is it to operate the simulation? • Can the centre provide independent documentation on acceptability? • Is the organisation/centre open to suggestions on content and new applications? • Will the simulation be considered relevant by your users and stakeholders (language, culture, local procedures, etc.)? • How difficult (time and money) is it to make the simulation relevant to your situation?
Learning	This defines specific changes in skills, knowledge, and attitudes in the target audience. Learning objectives are particularly applicable (but not exclusively) for training interventions • Is there any need for the program? • What are the performance objectives (task, conditions, and standards) the learners must master? • What specific aspects of the target procedure does the device simulate? • Does the simulator replicate the sensory inputs necessary to correctly perform the target procedure? If not, what is missing? • Has the simulation content been validated by recognised subject-matter experts? • Does the simulation provide performance feedback? • Has the simulation, when properly integrated with appropriate curricula, been proven to show transfer of learning when properly integrated and used with an approved curriculum? • Does the simulation have validated metrics? • Has the simulation been proven to distinguish between novice and expert performers? • Has the simulation been proven to have predictive validity? • Does the simulation have a user database that records performance? Can these data be exported? Is it secure? • How focused is the implementation Team on your area of clinical education?
Application and implementation	This defines the level of success for the intervention, often in terms of utilisation and sustainment over time. • Do you have implementation objectives or parameters (when, where, how many, how much)? • Can the centre demonstrate quality control in the production process? • Does the simulation equipment meet local government standards? • Will the simulation work in your local conditions (electrical, humidity, temperature, space requirements)? • What is the failure rate/malfunction of the simulation equipment? • Is there any infrastructure to control and operate the simulation equipment? • What type of warranty and/or service support does the equipment supplier offer? Is it convenient and timely for your location(s)? Can the centre provide proof of business stability and long-term sustainability? • Does the centre offer suitable payment models? • What are the centre's future plans for product additions and/or upgrades? • What are the licensing parameters and restrictions? Will you have a perpetual license to use the simulation or must the license be periodically renewed? • Can the simulator network with other simulators and learning management systems? • Is there any on-site or remote support from the equipment supplier?

(continued)

Table 3. Continued.

Objective	Comment
Impact	Also referred to as the business impact, this generally comes from the gap analysis and defines the specific impact or change expected from the intervention • Are there specific gaps, in either individual or organisational performance, you need to address (complication rates, learner time to proficiency, team performance etc.)? • Can the organisation provide you examples of how others have used the centre's simulation equipment to bridge similar performance gaps? • Is there any plan to measure the impact?
Return on investment	This defines the actual cost versus benefits expected from the intervention. It is generally defined as ROI (%) = Net benefits / Net costs X 100 • Do you have specific, financial ROI targets? • Can the organisation provide independently verified ROI case studies of situations similar to yours? • Does the organisation have an independently verified ROI model?
Intangibles benefits	This defines those effects which are not addressed by the other objectives above. • Can the organisation provide impartially verified examples of intangible benefits?

measured (Table 3), and the relevance of the 18 items towards trainer improvements during simulated teaching. This has both theoretical and practical importance, as its future use allows trainers to analyse and critically reflect on their abilities better than is possible with existing methods. This NTS-TeST addresses issues with NTS that some experts may have when performing or teaching. For example, they may rely on increased automated action and may not be able to appropriately deconstruct their cognitive processes into steps as compression of their thought and actions has occurred.[17] With a layout of 18 items that can be grouped into more specific categories, this NTS-TeST can more specifically identify the NTS which requires improvement from the trainers. It is suggested that this instrument can be used by both trainers and trainees to better quantify evaluation of a trainer's NTS performance.

Statistical analysis supports the theoretical justification for the inclusion of each question. For example, the questions asked before the start of training had good-to-strong reliability and validity and all had the objective to assess a trainers' ability to prime their student's mental models/schema before the start of training. It is through these homogenous questions that the trainer's aptitude to prepare their students can be measured. Preparation, expectations, and explanation of a simulation before starting can benefit the trainees' skills as they have been shown to both increase performance during training and aid in consistent improvements in performance.[18] However, Cronbach's alpha coefficient was 0.73 which was short of the desired =>0.8. The desired alpha would suggest robust internal consistency. One factor which affects alpha is sample size, specifically if there is a small sample size then large standard errors may reduce power. Increased sample size would reduce this issue if this tool were developed from this pilot/initial study. Each question was investigated for its effects on alpha if removed but none reached above 0.8, therefore this may be a combinatorial effect of the sample size with a possible latent variable in limiting the alpha produced. An exploratory factor analysis in subsequent development could discover any factor structure of the tool and be an addition to examine its internal reliability.

The questions used during the simulation were agreed upon to better understand the awareness and cognitive load of the trainees, and to better adjust their trainer's intervention behaviours if required. The feedback provided can modify trainers' pre-existing knowledge structures and consequences regarding how best to teach during simulation exercises.[7] This can aid in the prevention of demotivation and other negative outcomes of both parties. Trainers can be motivated to improve these metrics, and trainees are then motivated to improve their own training input. Approximately 43.1% of UK trainees stated they experienced negative events when in surgical training as there can be pressures, such as unrealistic knowledge expectations, leading to a poor learning experience.[19] Avoidance or reduction of such undesired occurrences with this NTS feedback is possible, promoting a crucial factor in the outcome of training success. Indeed, for both parties to gain positive experiences and increase the likelihood of trainees' skills acquisition, cognitive load also needed consideration.[20,21] Questions that directly or indirectly infer the fundamental load of the learning task (the scenario) and the extraneous load of the learning situation (the simulated anatomy and instruments) on trainees and trainers can allow adaptation to reduce excessive cognitive load in these areas. Cognitive load management has been increasingly used in medical simulation[22] but its utility in surgical technical skills simulation is lagging behind.

A single method to evaluate effective training by a trainer is inadequate as each method has its own limitations. Therefore, a triangulation process should be used to achieve a global view. We therefore plan to use the new NTS-TeST for self-evaluation by trainers as well as trainee ratings to facilitate an effective structured feedback to skills trainers. The NTS-TeST can be used as a "checklist" by trainers to remind them of important ingredients of an effective simulation-based skills teaching session. It is important for the course facilitator to consider how feedback should be delivered to the trainer to ensure that it does not have a negative effect.

The NTS-TeST presented here has the potential to be further developed. If data collection collaboration is possible then testing with large datasets and performing Confirmatory Factor Analysis can better assess validity. Implementation of this instrument can promote a culture that encourages both positive and negative feedback in a structured manner that promotes healthy, constructive, and safe improvements, and dissemination of these enhancements to other trainers. With the effects of COVID-19 disrupting face-to-face simulation training opportunities, online training sessions may benefit from a digital version of the NTS-TeST to be used with virtual reality simulation for trainees to help their trainers adapt to the new landscape.

Conclusion

Simulation-based education is progressively being used as an educational instrument in healthcare training programs. Becoming a successful trainer requires a variety of skills to foster an ability to create an environment for training that can incorporate a learner's unique experiences and learning goals. After modification to a previous instrument, this paper describes the development and assessments of efficacy of a novel NTS-TeST that has shown high internal consistency and reliability to measure trainers' NTS, as experienced by their trainees during technical skills simulation in surgery. This innovative NTS-TeST can be used alongside existing mechanisms to provide a holistic assessment of a trainer by their trainees.

Acknowledgements

We would like to thank all participants and faculty from Leeds and Belgium boot Camp for their support. We could not have done without the excellent help from Medical Education Team, Leeds Teaching Hospital, Leeds.

Informed consent

All participants in this study provided written informed consent.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Equipment and sponsorship for the Urology Simulation Boot Camp were provided by: Karl Storz, Cook Medical, Coloplast, Ethicon, Dantec, OKB Medical (Simbionix), MediPlus, Teleflex, European Pharma. M.P. and F.M. received from funding from an MRC Confidence in Concept Award (MC/PC/17,165) and F.M. was supported by a Turing Institute Fellowship.

ORCID iDs

Chandra Shekhar Biyani https://orcid.org/0000-0003-3443-8425
Marina Yiasemidou https://orcid.org/0000-0002-2599-4131

References

1. Taylor EW. Teaching Beliefs of Graduate Students in Adult Education: A Longitudinal Perspective, Adult Education Research Conference. https://newprairiepress.org/aerc/2002/papers/63 (2002, accessed 2 April 2021).
2. Pears M, Peebles D, Ward P, et al. Surgeons' nontechnical skills rapidly acquired by trainees – can human factors and virtual reality create synergistic and predictive effects? In: Gore J and Ward P (eds) *Naturalistic Decision Making and Uncertainty, Proceedings of the 13th bi-annual international conference on Naturalistic Decision Making*, The University of Bath, 2017, pp. 341–350.

3. Dean BJ, Keeler B, Garfjeld Roberts P, et al. Development of a surgical trainer assessment questionnaire. *ANZ J Surg* 2018; 88: 45–49.
4. Klein GA. *Strategies of decision making*. Yellow Springs, OH: Klein Associates Inc., 1989.
5. Ravindran S, Haycock A, Woolf K, et al. Development and impact of an endoscopic non-technical skills (ENTS) behavioural marker system. *BMJ Simul Technol Enhanc Learn* 2021; 7: 17–25. bmjstel-
6. Waschke KA, Anderson J, Macintosh D, et al. Training the gastrointestinal endoscopy trainer. *Best Pract Res Clin Gastroenterol* 2016; 30: 409–419.
7. Stice JE. Using Kolb's learning cycle to improve student learning. *Eng Educ* 1987; 77: 291–296.
8. Fleming S, Winson E, Amin K, et al. Evaluating trainee experience of surgical skills teaching. *J Surg Simul* 2015; 2: 1–5.
9. Biyani C, Hanchanale V, Rajpal S, et al. First urology simulation boot camp in the United Kingdom. *Afr J Urol* 2017; 23: 258–267.
10. Balasubramanian CK. The community balance and mobility scale alleviates the ceiling effects observed in the currently used gait and balance assessments for the community-dwelling older adults. *J Geriatr Phys Ther* 2015; 38: 78–89.
11. Howe J, Inness E, Venturini A, et al. The community balance and mobility scale-a balance measure for individuals with traumatic brain injury. *Clin Rehabil* 2006; 20: 885–895.
12. Hulley SB. *Designing clinical research*. Philadelphia: Lippincott Williams & Wilkins, 2007.
13. Tavakol M and Dennick R. Making sense of Cronbach's alpha. *Int J Med Educ* 2011; 2: 53–55.
14. Hilbert S, Kuechenhoff H, Sarubin N, et al. The influence of the response format in a personality questionnaire: an analysis of a dichotomous, a Likert-type, and a visual analogue scale. *TPM: Test Psychometr Methodol Appl Psychol* 2016; 23: 3–24.
15. Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd ed. Mahwah, NJ: L. Erlbaum Associates, 1988.
16. George D and Mallery P. *IBM SPSS statistics 26 step by step: a simple guide and reference*. Abingdon: Routledge, 2019.
17. Moulton C, Regehr G, Lingard L, et al. Slowing down to stay out of trouble in the operating room: Remaining attentive in automaticity. *Acad Med* 2010; 85: 1571–1577.
18. Lipshitz, R, Shaul B and Schemata O Mental Models in Recognition-Primed Decision-Making, In: Zsambok C and Klein G (eds) *Naturalistic decision-making, Lawrence Erlbaum Associates*, Mahwah, NJ, 1997, pp. 293–303.
19. Chapman SJ, Hakeem AR, Marangoni G, et al. How can we enhance undergraduate medical training in the operating room? A survey of student attitudes and opinions. *J Surg Educ* 2013; 70: 326–333.
20. Miller GA. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol Rev* 1956; 63: 81–97.
21. Sweller J. Cognitive load during problem solving: effects on learning. *Cogn Sci* 1988; 12: 257–285.
22. Andersen SAW, Mikkelsen PT, Konge L, et al. Cognitive load in distributed and massed practice in virtual reality mastoidectomy simulation. *Laryngoscope* 2016; 126: E74–E79. doi:10.1002/lary.25449