Psychometric Testing of a Structured Assessment Instrument for Non-technical Skills (NANTS-no) for Use in Clinical Supervision of Student Nurse Anesthetists

Flynn, Fiona Mary[1]; Valeberg, Berit Taraldsen[2]; Tønnessen, Siri[1]; Bing-Jonsson, Pia Cecilie[3]

[1]Institutt for sykepleie- og helsevitenskap - Universitetet i Sørøst-Norge
[2]Institutt for sykepleie og helsefremmende arbeid - OsloMet - storbyuniversitetet
[3]Fakultet for helse- og sosialvitenskap - Universitetet i Sørøst-Norge

This is a post-peer-review, pre-copyedit version of an article published in *Journal of Nursing Measurement*. The final authenticated version is available online at: http://dx.doi.org/10.1891/JNM-D-19-00086

Publisher's version: Flynn, F. M., Valeberg, B. T., Tønnessen, S., & Bing-Jonsson, P. C. (2020). Psychometric Testing of a Structured Assessment Instrument for Non-technical Skills (NANTS-no) for Use in Clinical Supervision of Student Nurse Anesthetists. *Journal of Nursing Measurement*, *29*(1), 1-20. https://doi.org/10.1891/JNM-D-19-00086

Title

Psychometric testing of a structured assessment instrument for non-technical skills (NANTS-no) for use in clinical supervision of student nurse anaesthetists

Authors

Fiona M. Flynn*, CRNA, MSc
[1]Department of Nursing and Health Sciences, Faculty of Health and Social Sciences, University of South-Eastern Norway (USN), Postbox 235, 3603 Kongsberg, Norway.
*Corresponding author: Telephone: +47 975 31 697, e-mail: fiona.flynn@usn.no

Berit T. Valeberg, CRNA, Ph.D, Professor
[1]Department of Nursing and Health Sciences, Faculty of Health and Social Sciences, University of South-Eastern Norway (USN)
[2]Department of Nursing and Health Promotion, Faculty of Health Sciences, Oslo Metropolitan University

Siri Tønnessen, MSc, Ph.D
[1]Department of Nursing and Health Sciences, Faculty of Health and Social Sciences, University of South-Eastern Norway (USN)

Pia Cecilie Bing-Jonsson, MSc, PhD
[1]Department of Nursing and Health Sciences, Faculty of Health and Social Sciences, University of South-Eastern Norway (USN)

Acknowledgements

The authors would like to thank the nurse anaesthetists, who took part in the study and the hospital trusts who made participation possible. A special thanks to Kjersti Sandaker, Henny-Mari Devold Hagen, Egil Bekkhus and Kjersti Forbech Henriksen for their help with assessing the films, to Knut A. Hagtvet for his invaluable help on generalizability theory, and Ragnhild Sørum Falk for her help with the statistical analyses.

Conflict of interests

The authors declare they have no competing interests.

Funding

This project has received no financial support.

ABSTRACT

Background:

This study evaluated psychometric properties of a structured behavioural assessment instrument, Nurse Anaesthetists' Non-Technical Skills- Norway (NANTS-no). It estimated whether reliable assessments of non-technical skills could be made after taking part in a workshop. An additional objective was to evaluate the instrument's acceptability and usability.

Methods:

An explorative design was used. Nurse anaesthetists (n=46) involved in clinical supervision attended a six-hour workshop on non-technical skills, then rated non-technical skills in video-recorded simulated scenarios and completed a questionnaire.

Results:

High reliability and dependability were estimated in this setting. Participants regarded the instrument as useful for clinical supervision of student nurse anaesthetists.

Conclusions:

Findings suggest that NANTS-no may be reliable for performing clinical assessments of student nurse anaesthetists and encouraging critical reflection. However, further research is needed to explore its use in clinical settings.

Keywords

Non-technical skills, Nurse anaesthetist/anesthetist, Psychometric testing, Patient safety, Education, NANTS-no

Introduction

Clinical supervision is an essential part of anaesthesia nursing education. It strives to bridge the gap between theory and practice and promote professional and personal development, by facilitating the learning of clinical skills and developing critical and reflective thinking (Jokelainen, Turunen, Tossavainen, Jamookeeah, & Coco, 2011). However, the majority of Norwegian nurse anaesthetists involved in clinical supervision currently lack any formal training in this area. A recent report on the quality of clinical practice (The Norwegian Association of Higher Education Institutions, 2016) called for closer cooperation between educational and healthcare institutions, as well as a more formalized training for clinical supervisors to raise standards.

The International Federation of Nurse Anesthetists (IFNA, 2016) promotes high standards of competence and behaviour to ensure quality in anaesthesia. Developing, training and assessing non-technical skills (NTS) is generally regarded as essential for providing safe anaesthesia and ensuring excellent care (Fletcher et al., 2003; Flin & Mitchell, 2009; Glavin, 2009). However, there is a need for a common taxonomy and robust and reliable instruments in anaesthesia nursing education for observing and assessing these skills. The Nurse Anaesthetists' Non-Technical Skills – Norway (NANTS-no) structured behavioural assessment instrument is intended for developing and assessing non-technical skills in student and postgraduate nurse anaesthetists in Norway, and was tested in a simulation setting (Flynn, Sandaker, & Ballangrud, 2017). In this paper, the instrument's psychometric properties were further tested, prior to using the instrument in clinical settings. The purpose of the study was to:

- Explore whether experienced nurse anaesthetists involved in clinical supervision can reliably and accurately assess NTS in simulated video-recorded scenarios using NANTS-no, after participating in a six-hour workshop.
- Estimate whether each individual mentor is able to provide a reliable assessment of NTS in video-recorded simulated scenarios.
- Explore whether NANTS-no is perceived as an acceptable and usable instrument for developing and assessing student nurse anaesthetists' NTS in clinical practice.

Background and conceptual framework

Clinical supervision is a generic term for the process of providing support and guidance to student nurse anaesthetists (SNAs) or other professionals, with the aim of enabling learning and development of professional skills in a safe environment (Lyth, 2000). Proctor (1991) described clinical supervision as having normative, formative and restorative elements, where the *normative* focuses on the setting, the *formative* on education and development, and the *restorative* on providing support. The term *mentor* appears to have a variety of meanings in clinical practice literature, encompassing both ad-hoc arrangements and more formalized monitoring and assessment (Fowler & Cutcliffe, 2011; Jokelainen et al., 2011). In this paper, mentor is used to describe the experienced postgraduate nurse anaesthetists whose role involves supervising, teaching, and assessing SNAs throughout their clinical training.

Mentorship places responsibility on both the mentor and the SNA to enable an individual learning process and empower development of a new professional identity and competence (Jokelainen et al., 2011). Although there are standardized means of testing theoretical knowledge, there is currently a lack of validated and reliable instruments for forming and assessing nurse anaesthetists in clinical practice in Norway, and few internationally (Collins & Callahan, 2014; Lyk-Jensen et al., 2016; Sevdalis, Hull, & Birnbach, 2012). A major challenge facing mentors is simultaneously ensuring patient safety, while guiding SNAs through complex, dynamic and critical situations in a highly technical environment. A recent study highlighted the way in which management attitudes and increasing demands for efficient production limit professional development among nurse anaesthetists (Averlid, 2017). These same factors are among those reported by both students and mentors as constraints on the mentoring role and a threat to patient safety (Jølstad, Røsnæs, Lyberg, & Severinsson, 2017; Rylance, Barrett, Sixsmith, & Ward, 2017).

The IFNA "Code of Ethics and Standards of Practice, Monitoring and Education" for nurse anaesthetists utilizes the Canadian Medical Education Directions for Specialists (CanMEDs) Competency Framework (IFNA, 2016). The CanMEDs framework has been recently adopted by the Norwegian Association of Nurse Anesthetists in an attempt to ensure high standards of safety and quality in clinical practice and the education of nurse anaesthetists. The seven competencies described in the framework (Figure 1) incorporate non-technical skills such as communication and situation awareness, task management, leadership and teamwork, as well as promoting the professional identity of the nurse anaesthetist (Herion, Egger, Greif, & Violato, 2019). The decision to adopt this framework for Norwegian nurse anaesthetists is in line with an international movement in health care education and clinical practice aimed at moving beyond competence, with excellence as an aspirational goal (O'Donnell, Cook, & Black, 2016; Smith, Glavin, & Greaves, 2011; Wong, 2012). Since the role of non-technical skills in developing standards of clinical excellence and improving patient safety is the focus of this paper, the CanMEDs framework seemed an appropriate conceptual framework for the measurement of non-technical skills.

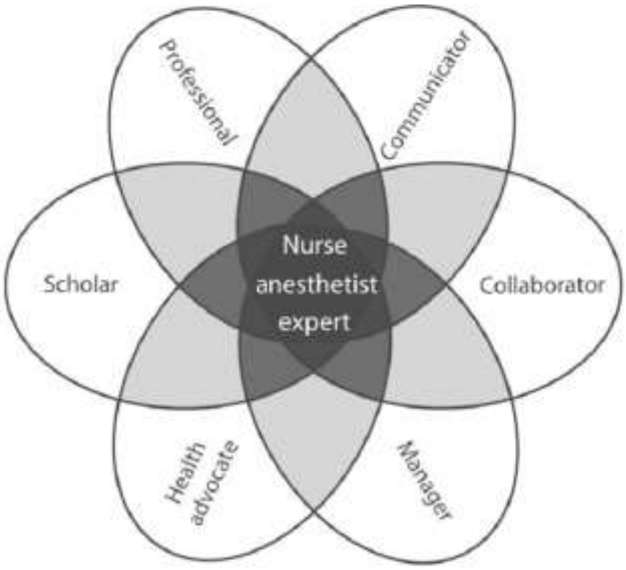

Figure 1: Adapted by IFNA from the CanMEDS Physician Competency Diagram with permission from The Royal College of Physicians and Surgeons of Canada. Copyright©2009

There is a general acceptance that a high number of adverse events in healthcare are a result of human factors and could have been prevented (Jha, Prasopa-Plaizier, Larizgoitia, & Bates, 2010). Surgery and anaesthesia are particularly high-risk areas (Kennerly et al., 2014), where simple mistakes can have fatal consequences. Measures have been implemented at all levels to improve patient safety in the operating theatre (Haugen et al., 2017; Sevdalis et al., 2012; Tan, Pena, Altree, & Maddern, 2014), and there is a growing bank of research on the role of NTS that focuses on the individual professional's behaviour. Flin et al. (2008) define NTS as "cognitive, social and personal resource skills that complement technical skill, and contribute to safe and efficient task performance". Ensuring nurse anaesthetists have well-developed NTS, such as heightened situation awareness, optimal decision-making and task management, and open and effective communication is a prerequisite for excellent practice and providing high standards of safe anaesthesia care (Gaba, Howard, & Small, 1995; Herion et al., 2019; Larsson & Holmstrom, 2013; Rutherford, Flin, & Mitchell, 2012).

In response to international focus on the role of human factors in adverse events, a number of behavioural assessment instruments have been developed to structure the development, training and assessment of NTS in healthcare professionals (Higham et al., 2019). There are currently instruments for assessing NTS in an operating theatre setting for anaesthesiologists (Fletcher et al., 2003), nurse anaesthetists (Lyk-Jensen, Jepsen, Spanager, Dieckmann, & Ostergaard, 2014), anaesthetic practitioners (Rutherford, Flin, Irwin, & McFadyen, 2015), surgeons (Spanager et al., 2013; Yule, Flin, Paterson-Brown, Maran, & Rowley, 2006) and scrub nurses (Mitchell et al., 2013). NANTS-no was adapted for Norwegian nurse anaesthetists in 2014 (Flynn et al., 2017) from Anaesthetists' Non-Technical Skills (ANTS) for anaesthesiologists (Fletcher et al., 2003). The structure and sequence of the categories and the decision to change the rating scale from four to five points, was influenced by the Danish instrument for nurse anaesthetists (Lyk-Jensen et al., 2014). ANTS was chosen as the basis for NANTS-no as it had already been translated to Norwegian and seemed most appropriate, requiring only relatively minor changes to the language and practice examples to make it acceptable. It had also already been validated in 2014, unlike the Danish instrument (Fletcher et al., 2003; Lyk-Jensen et al., 2016).

NANTS-no has a hierarchical structure of four categories and fifteen elements (Figure 2), with behavioural examples of good and poor practice for each element. In addition, the instrument has a five-point numerical rating scale (1-5), where behaviour that places the patient's life at risk is rated as 1, marginal behaviour as 2, acceptable behaviour as 3, good behaviour as 4 and excellent behaviour as 5. The numerical rating scale is used to rate each element and category, and to provide a global score. Where behaviour is not observed for an element, N is used.

Categories	Category score	Elements	Element score
Situation awareness		Gathering information	
		Recognizing and understanding	
		Anticipating and thinking ahead	
Decision making		Identifying possible options	
		Assessing risks and selecting options	
		Re-evaluating	
Task management		Planning and preparing	
		Prioritizing	
		Identifying and utilizing resources	
		Maintaining standards and levels of quality	
Team working		Exchanging information	
		Assessing roles and capabilities	
		Co-ordinating activities	
		Displaying authority and assertiveness	
		Supporting other team members	

Figure 2: The NANTS-no framework

All these behavioural rating instruments have a similar taxonomy and rating scales, and are designed to facilitate objective observations and ratings of the skills and behaviour relevant for the individual professional (Flin & Maran, 2015; Higham et al., 2019). The purpose of NANTS-no and other similar instruments is primarily formative enabling mentors to provide structured and objective feedback (Jepsen, Østergaard, & Dieckmann, 2015). A further aim is to encourage self-awareness and critical thinking, allowing healthcare professionals to reflect on and assess how their behaviour affects their performance and highlighting issues that need addressing. Assessment is important both to ensure quality and validity in education, encourage learning and demonstrate accountability to stakeholders (Wong, 2012). Traditionally, assessment in anaesthesia nursing education has focused more on testing theoretical knowledge and technical proficiency. In order to be able to make systematic summative assessments of NTS in clinical practice, the instruments must be sufficiently reliable and robust (Higham et al., 2019). Regular use and repeated training are recommended before using these instruments to make assessments in clinical settings (Flin & Maran, 2015).

The need for separate instruments that reflect the healthcare setting in which they are to be used has been discussed (Flin & Patey, 2011; Higham et al., 2019; Jepsen et al., 2015; Wisborg & Manser, 2014). Since a generic instrument for assessing NTS in healthcare does not currently exist, there is a persuasive argument for a separate instrument that reflects organizational and cultural differences in anaesthesia nursing education and clinical practice. Nurse anaesthetists in Norway have an independent professional responsibility when providing anaesthesia, as well as working in close collaboration with anaesthesiologists (Averlid, 2017; Nilsson & Jaensson, 2016; Ringvold et al., 2018), and NANTS-no was adapted to reflect this setting. Although NANTS-no appeared to display high reliability (ICC = 0.91) when used in a simulation setting, there is need for further research to see whether it may also be suitable for use in clinical settings.

Methods

The development of NANTS-no has been described in an earlier paper (Flynn et al., 2017). The current study used an explorative design to examine the instrument's psychometric properties prior to use in clinical settings. Participants used the NANTS-no five-point rating scale to rate NTS in video-recorded simulated scenarios. These ratings together with an evaluation questionnaire provided the data for testing the instrument's reliability, generalizability, acceptability and usability.

Sample

A convenience sample of 69 nurse anaesthetists involved in clinical supervision of SNAs in four hospital trusts in Norway, were invited to take part in a six-hour workshop on NTS in anaesthesia nursing. Forty-six nurse anaesthetists consented to take part in the study, which took place over a four-week period in October and November 2017. An additional participant, who attended the workshop, was excluded from the study owing to having been involved in the production of the video-recorded simulated scenarios. All the participants were actively involved in clinical supervision of SNAs, either as mentors (*n* = 35), clinical supervisors (*n* = 3) or with responsibility for professional development (*n* = 8). Background information relating to participants' sex, age and experience was collected (Table 1).

Preparatory phase

Video clips

Six short video-recorded simulated scenarios were specially designed and produced for the study, showing nurse anaesthetists displaying varying levels of NTS. The scenarios for the video clips were developed by the main researcher (F.M.F) and a team of experienced nurse anaesthetists involved in simulation training of NTS. They were evaluated for content validity by the nurse anaesthetists who took part in filming the scenarios. Having read the scripts, they provided feedback so changes could be made. The scenarios were loosely scripted to ensure that they demonstrated the desired NTS, while enabling a dialog that was as natural as possible. They featured student and expert nurse anaesthetists, as well as other members of the anaesthesia and surgical team in a variety of perioperative situations. These included critical and routine situations that can occur during intubation and extubation of anaesthetized patients as well as before and during surgery. Each video clip lasted between four and eight minutes. They were filmed using a Laerdal Medical® patient simulator, with trained nurse anaesthetists and other members of the surgical team playing different roles.

Reference ratings

A set of reference ratings for the video clips used in the study were produced by a panel of four experts, all with relevant clinical experience and considerable expertise and interest in educating and training NTS (Keeney, Hasson, & McKenna, 2011). Two members of the panel were involved in the development and testing of NANTS-no in a previous study (Flynn et al., 2017), while a third member had experience in using NANTS-no for educating SNAs. The fourth member had considerable expertise in educating and training NTS with critical care nurses and other groups of healthcare professionals. Each member of the panel rated the video clips alone prior to a meeting, where any disparity in the ratings was discussed face-to-face in order to reach a consensus (Keeney et al., 2011).

Evaluation questionnaire

An evaluation questionnaire based on one used in other studies (Mitchell et al., 2013; Rutherford et al., 2015) was translated to Norwegian with permission from John Rutherford, and then adapted. The translation was carried out by the main researcher who is bilingual in both English and Norwegian, and a colleague. The evaluation questionnaire was used to collect background information, evaluate the workshop and assess the acceptability and usability of NANTS-no for use in clinical supervision.

Setting

The workshop was held by the main researcher on seven different occasions either at the various hospital trusts or at a nearby venue, in order to encourage participation. Prior to taking part in the workshop, the participants were asked to familiarize themselves with NANTS-no and watch a film on the role of human factors in anaesthesia (*Just a routine operation*: https://www.youtube.com/watch?v=VndU2zap_Rg). The workshop comprised theory about patient safety, human factors and the underlying concepts for developing and assessing NTS, as well as group discussion and rater training. The structure of NANTS-no, its categories and elements were discussed in detail. Participants were encouraged to give examples of both good and poor NTS, either that they had experienced personally or witnessed while working with SNAs. The rating scale and appropriate use of all the scores (1-5) as well as N (not observed) was also explained. Rater training involved rating three training video clips, after which ratings were discussed, and the participants given feedback. The training video clips were not used for data collection in the study.

At the end of the workshop, each participant individually rated NTS in the six video clips produced for the study using the NANTS-no rating scale. NTS were rated at element and global level. Participants were also asked to complete the evaluation questionnaire.

Psychometric testing of the assessment instrument

Reliability

Reliability of ratings is relative to the proportion of systematic and random variance inherent in the measurements. Systematic variance is the true differences between the nurse anaesthetists rated in the video clips and random variance is the error component present in the actual ratings (Streiner, Norman, & Cairney, 2015). Reliability was assessed using analyses based on classical test theory; internal consistency, inter-rater reliability and rater accuracy.

Internal consistency was estimated for each category across all the videos using Cronbach's alpha. Inter-rater reliability was estimated using a two-way mixed, absolute agreement Intra-class Correlation Coefficient (ICC), where an ICC value > 0.75 represents high reliability, ICC 0.5 - 0.75 represents moderate reliability and ICC < 0.5 poor reliability (Portney & Watkins, 2009). As the study had a large number of raters in relation to the number of NANTS-no elements and is intended for use in clinical practice with one or two raters, an ICC based on many raters could be misleading (Streiner et al., 2015). Inter-rater reliability was thus measured at all levels using a mean ICC derived from five randomly selected pairs of raters.

The accuracy of the participants' ratings were compared to the set of reference ratings. A system of points was assigned to reflect how far the participants' ratings deviated from the reference ratings. Participant ratings that were the same as the reference ratings were assigned 5 points. A one-point deviation was assigned 4 points, while a two-point deviation was assigned 3 points and so on. The raters' total score was then calculated and presented as a percentage of the expert total score for each NANTS-no element across all the video clips. Since some raters had used two scores such as 2-3, rater accuracy was also assessed for one scale point difference from the reference score. The mean absolute deviation (MAD) from the reference ratings was calculated.

Generalizability

Generalizability theory (GT) was used to estimate the various error components and explore the dependability of the ratings for future use (Brennan, 2010). A generalizability (G) study with a balanced two-facet crossed design was carried out with video clips (n = 6) x raters (n = 46) x NANTS-no categories (n = 4). The estimated variance components from the G-study were then used to estimate an absolute generalizability coefficient for the number of raters needed to reliably rate the films in a decision (D) study, where the categories were a fixed component.

Acceptability and usability

To assess whether NANTS-no was able to measure different types of behaviour, the instrument was tested for observability, acceptability and usability using data from the evaluation questionnaire. The level of observability of NTS in the videos was also assessed by comparing the percentage of "not observed" with the "observed" scores for each video clip.

Data analysis

Data was analysed using IBM SPSS Statistics, version 24, and the generalizability analyses were performed using the MATLAB G1.sps program for SPSS (Mushquash & O'Connor, 2006). A random number generator in Microsoft Excel was used for selecting pairs of raters for the inter-rater reliability analysis.

Any NANTS-no ratings that were given as two scores, for example 2-3, were rounded down to the lower score, while "not observed" was treated as zero. Category scores were calculated as the mean score of the elements in each category. A missing data analysis was carried out to ensure that it was less than 3%, and any missing data was replaced with zero in the generalizability analyses.

Ethical considerations

Following notification to the Norwegian Centre for Research Data (project no. 55538) on 8.9.2017 and approval from the hospital trusts, the nurse anaesthetists were informed orally and in writing about the study. In accordance with the Declaration of Helsinki (World Medical Association, 2013), the concepts of informed consent, voluntary participation and the right to withdraw without penalty were carefully explained to the participants. Requirements regarding confidentiality, data anonymity and secure handling of data were also explained. After appropriate time for consideration, written consent was obtained from all participants.

RESULTS

Forty-six nurse anaesthetists involved in clinical supervision of SNAs took part in the workshops. Two participants rated only four out of six clips and a third participant rated five of the clips. Forty-one participants completed the evaluation questionnaire at the end of the workshop. The characteristics of the sample are presented in Table 1.

Table 1: Characteristics of the sample

	%	N	Min	Max	Mean (SD)
Sex:		46			
Male	19.6				
Female	80.4				
Age in years:		46	31	62	47.3 (8.2)
Number of years as nurse anaesthetist:		46	1	30	12.5 (7.4)
Hospital Trust 1	34.8	16			
Hospital Trust 2	23.9	11			
Hospital Trust 3	30.4	14			
Hospital Trust 4	10.9	5			
Previous experience with clinical supervision of SNAs		41			
Yes	73.2				
No	26.8				
Some previous experience with NANTS		41			
Yes	46.3				
No	53.7				

Reliability

A total Cronbach's α >0.9 was estimated in all categories. The overall inter-rater reliability was estimated as high (ICC = 0.80), and moderate to high for the NANTS-no elements (mean ICC = 0.68-0.91). Inter-rater reliability and Cronbach's α at category level is presented in Table 2.

Table 2: Inter-rater reliability (ICC) for five randomly selected pairs of raters and Cronbach's alpha

NANTS-no category	Intra-class Correlation Coefficient (95% CI) pair 1	pair 2	pair 3	pair 4	pair 5	Mean ICC	Cronbach's alpha
Situation awareness	0.76 (0.22-0.84)	0.83 (0.19-0.95)	0.80 (0.40-0.93)	0.90 (0.74-0.96)	0.81 (0.48-0.93)	0.82 (0.77-0.87)	0.95
Decision making	0.85 (0.61-0.95)	0.46 (-0.21-0.78)	0.80 (0.44-0.93)	0.92 (0.78-0.97)	0.90 (0.68-0.96)	0.79 (0.62-0.95)	0.94
Task management	0.67 (0.25-0.85)	0.79 (0.33-0.92)	0.72 (0.35-0.88)	0.93 (0.85-0.97)	0.81 (0.57-0.92)	0.78 (0.70-0.87)	0.91
Team working	0.73 (0.28-0.88)	0.62 (0.14-0.83)	0.80 (0.33-0.92)	0.94 (0.87-0.97)	0.86 (0.70-0.99)	0.79 (0.68-0.90)	0.93
Global score	0.81 (-0.64-0.98)	0.79 (-0.13-0.97)	0.87 (0.26-0.98)	0.90 (0.24-0.99)	0.95 (0.65-0.99)	0.86 (0.81-0.92)	
Total inter-rater reliability	0.74 (0.60-0.83)	0.68 (0.27-0.84)	0.78 (0.52-0.88)	0.92 (0.88-0.95)	0.85 (0.77-0.90)	0.80 (0.71-0.88)	

A mean rater accuracy of 82% of the maximum expert element score was estimated. Mean rater accuracy increased to 89% of the expert score when estimated to one scale point difference. Rater accuracy for global scores was estimated as 81% (MAD = 1.08) and 87% (MAD = 1.19) to one scale point difference. Mean percentages and mean absolute deviation from the total expert score for the individual elements are presented in Table 3.

Table 3: Rater accuracy at element level

NANTS-no elements	Mean % of expert ratings (MAD)	Mean % ratings accurate ±1 scale point (MAD)
Gathering information	83 (0.78)	91 (0.86)
Recognizing and understanding	88 (0.74)	95 (0.58)
Anticipating and thinking ahead	87 (0.78)	94 (0.68)
Identifying possible options	86 (0.71)	96 (0.60)
Assessing risks and selecting options	83 (0.96)	89 (1.05)
Re-evaluating	86 (0.89)	92 (0.85)
Planning and preparing	86 (0.68)	96 (0.55)
Prioritizing	87 (0.79)	94 (0.68)
Identifying and utilizing resources	83 (0.89)	91 (0.93)
Maintaining standards and levels of quality	77 (1.25)	84 (1.40)
Exchanging information	81 (0.87)	89 (1.02)
Assessing roles and capabilities	80 (0.85)	91 (0.98)
Co-ordinating activities	76 (1.27)	83 (1.43)
Displaying authority and assertiveness	79 (1.04)	87 (1.18)
Supporting other team members	65 (1.87)	70 (2.03)

Generalizability

The G-study estimated an absolute error variance $\sigma^2_\Delta = 0.015$, with a higher degree of variance among the raters ($\sigma^2_r = 0.084$) than in the NANTS-no categories ($\sigma^2_c = 0.016$). There was a certain amount of variance in the raters scoring for each video clip ($\sigma^2_{pr} = 0.237$), but only minimal variance in the raters average level in the NANTS-no categories ($\sigma^2_{rc} = 0.007$). The D-study estimated that one rater was sufficient to achieve an absolute generalizability coefficient = 0.83 using NANTS-no (Figure 3).

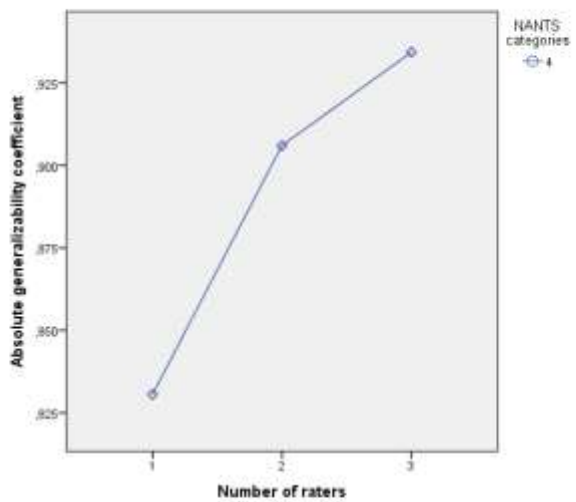

Figure 3: Effect of number of raters on the generalizability coefficient

Observability

Observability of NTS across the scenarios was high, averaging 91% (82.4 - 97.8%). However, 18 participants rated over 10% of the scores as "not observed". The distribution of "not observed" scores in the video clips is shown in Table 4. Although there was a high level of observability at category level, ranging between 83% and 95%, the following five NANTS-no elements had over 10 % of scores recorded as "not observed": *Re-evaluating* 15.9%; *Maintaining standards and levels of quality* 14.9%; *Assessing roles and capabilities* 15.9%; *Coordinating activities* 13.8%; and *Supporting other team members* 33%.

Table 4: Distribution of "not observed" scores in the video clips

Video	N	Mean %	Std. deviation (95% CI)
1	46	10.61	13.63 (6.56 - 14.66)
2	46	12.22	14.90 (7.79 - 16.64)
3	46	7.30	9.94 (4.35 -10.26)
4	46	2.22	4.91 (0.76 – 3.67)
5	46	16.85	17.39 (11.68 – 22.01)
6	46	6.52	8.52 (3.99 – 9.05)
Total	276	9.29	13.03 (7.74 – 10.83)

One-way ANOVA: F (5, 270) = 7.87, p = <0.005

Acceptability and usability

In response to the evaluation questionnaire, 97.5% of participants stated that NANTS-no described the NTS essential for a nurse anaesthetist either well or very well. NANTS-no was also described as a useful instrument for aiding SNAs to develop these skills (100%), promoting critical reflection (97.6%), providing feedback (97.6%) and evaluating SNAs in clinical practice (92.7%). Fifty-six percent of the participants responded that they were able to identify the NTS in the videos either well or very well, and 83% stated that they had received sufficient training about NANTS-no and the underlying concepts (95%) during the training session.

DISCUSSION

The NANTS-no structured behavioural assessment instrument aims to facilitate critical reflection, and development and assessment of NTS in clinical practice. The findings in this explorative study suggest that NANTS has sufficiently high reliability and dependability when rating NTS in video clips to encourage testing it in clinical practice. The nurse anaesthetists participating in the study also supported this view.

Reliable and accurate assessment of NTS

Although overall reliability in the ratings of the video clips was slightly lower than in a previous study that estimated an ICC = 0.91 (Flynn et al., 2017), both inter-rater reliability and internal consistency were higher at category level in this study. An internal consistency $\alpha > 0.9$ may imply that some of the elements are redundant, though such a high alpha is not unusual and may also be a result of factors such as the study setting or category diversity in the scale (Streiner et al., 2015). The overall reliability of ICC = 0.80 compares favourably with a similar study (Rutherford et al., 2015), though the Danish instruments for anaesthesiologists and nurse anaesthetists demonstrated higher levels of inter-rater reliability (Jepsen et al., 2016; Lyk-Jensen et al., 2016). Inter-rater agreement levels estimated for ANTS and SPLINTS in similarly designed studies were only reported as acceptable (Fletcher et al., 2003; Mitchell et al., 2013). The high reliability estimated in this setting is therefore encouraging.

The CanMEDs competency framework incorporates NTS in its seven competencies. While there is currently no gold standard for observing and assessing these skills (Higham et al., 2019), an evidence-based framework for developing behavioural assessment instruments and training instructors and those assessing NTS does exist (Hull et al., 2013; Klampfer et al., 2001). Rater training and familiarity with the structured assessment instrument is seen as essential prior to using these types of instruments in clinical practice (Flin & Maran, 2015). A two-day training program is recommended for this purpose (Hull et al., 2013; Klampfer et al., 2001). Despite only six hours training and over 50 % of participants having no previous experience with NANTS, reliability at category level was estimated as >0.75. Inter-rater reliability at element level was more variable (0.68 – 0.91) and not as high as for N-ANTS-dk (Lyk-Jensen et al., 2016), with the NANTS element *Maintaining standards and levels of quality* showing poorest reliability. In addition, only moderate reliability was estimated for the elements *Assessing risks and selecting options* (0.73*), Assessing roles and capabilities* (0.70) and *Supporting other team members* (0.72). This is still generally higher than in some other studies (Fletcher et al., 2003; Rutherford et al., 2015). *Assessing risks and selecting options* is a cognitive element and may be more difficult to observe (Fletcher et al., 2003), although the cognitive elements in the Situation awareness category were all estimated to have an ICC > 0.8. A possible reason for the lower reliability in the other three elements lies in the fact that more than 10% of the raters had scored these elements as "not observed", with a total of 33% scoring the element *Supporting other team members* as "not observed". Since "not observed" behaviour was represented in the data as zero, this inevitably had a negative effect on the analysis.

Rater accuracy reflects the participants' ability to distinguish good from poor behaviour (Sevdalis et al., 2012) which is an important factor when supervising SNAs in situations where small mistakes can dramatically affect patient safety. A mean rater accuracy of 82% for elements across all the video clips is comparable with other similar studies (Jepsen et al., 2016; Lyk-Jensen et al., 2016), though

ANTS reported higher rater accuracy (Fletcher et al., 2003). Since some of the raters were unable to decide on one score and rated some elements or global scores with for example 2-3, it seemed advisable to test rater accuracy to one point difference. This increased the mean rater accuracy. Although time had been spent on calibrating scoring in the workshop, there were presumably difficulties here as in other studies in deciding where to set the boundaries for each score (Fletcher et al., 2003), underlining the need for repeated calibration training. The variation in mean accuracy for the various elements (70 – 96%) is presumably also accounted for by the relatively high number of “not observed” scores for some elements, reducing total scores. Participants were encouraged to use the entire scale including “not observed”. However, there was an unanticipated higher use of “not observed” scores compared with the expert panel, which may have been a result of misunderstandings regarding when it should be used. Again surprisingly, rater accuracy is highest in the category Situation Awareness, despite cognitive behavioural skills being usually regarded as difficult to observe (Flin et al., 2008).

Dependability of individual ratings

Although the ICC>0.8 estimated in this study is consistent with the required reliability for using NANTS-no as a formative instrument (Hull et al., 2013), reliability testing does not provide information about the variation between the individual participants, and whether each individual participant is able to perform a reliable assessment. Since the aim of this study was to explore whether this structured assessment instrument has a future in formative and summative assessments in anaesthesia nursing education, a high generalizability coefficient is of paramount interest.

The G-study estimated that the greatest variance lies between the raters’ scores for each video clip, which is unsurprising as each video clip was designed to demonstrate varying levels of NTS. However, only minimal variance was estimated between the elements in each category, and the way the raters scored each category. Although this suggests that the video clips may not have displayed enough variation in the non-technical skill elements and categories in each clip, it may also be a result of difficulties in observing and placing different types of behaviour correctly (Graham, Hocking, & Giles, 2010; Rutherford et al., 2015). Lyk-Jensen et al. (2016) suggested that having to rate multiple video clips over a short space of time might prove wearisome (rater fatigue) and thus affect results. To a certain extent, the lack of variance is unsurprising, as poor behaviour in one element/category may easily have an impact on behaviour in another element/category. For example, poor situation awareness will affect decision making and prioritizing of tasks, while poor planning may affect team working and situation awareness.

The D-study estimated that one rater could achieve reliable ratings with a generalizability coefficient >0.8, which is acceptable for formative assessments. According to Spanager et.al. (2013), a generalizability coefficient >0.9 is recommended for high-stakes assessments for certification purposes. Since there is a need for reliable instruments that can be used for summative assessments to ensure SNAs achieve expected levels of competency as described in the CanMEDs framework, it is encouraging that the D-study estimated that two raters could achieve a generalizability coefficient of 0.91. However, these results were attained rating scripted simulated scenarios. Using a structured assessment instrument to observe behaviour of SNAs in clinical settings over longer periods while simultaneously ensuring patient safety, is recognized as far more demanding (Flin, Patey, Glavin, & Maran, 2010).

Acceptability and usability of NANTS-no

An important factor for the future integration of NANTS-no in anaesthesia nursing education is acceptability (Fletcher et al., 2003). Another factor is how easy the instrument is to use. As the aim is to provide a common taxonomy that SNAs, mentors and other educators can all use to express what is considered excellent anaesthesia care, it is positive that NANTS-no is regarded as describing the NTS essential in a nurse anaesthetist. The participants also regarded NANTS-no as useful for encouraging SNAs to reflect critically on their performance, as well as aiding mentors in structuring their feedback and providing formative assessments. Objective and specific feedback is particularly necessary when guiding students who are struggling to provide safe anaesthesia care (Flin & Maran, 2015). The high percentage of participants that supported this view suggests that NANTS-no would be acceptable as an instrument for this purpose.

Although 56% of the participants stated they were able to identify the NTS in the video clips well or very well, there was a particularly high use of “not observed” scores in the Team working category. This is surprising as social and inter-personal skills are usually easier to observe than cognitive skills (Flin et al., 2008). The two video clips with the highest mean percentage of “not observed” scores were the two where the expert panel had also used “not observed” for several elements. Even though time was spent on explaining and calibrating the different scores and the use of “not observed”, the relative high use of “not observed” for certain elements may have been due to misunderstandings, difficulties in differentiating elements or rater fatigue. Despite this, the overall reliability of the instrument was still estimated as high. A follow-up workshop to clarify these issues would be beneficial prior to using NANTS-no in clinical situations.

Strengths and limitations

This study has a similar design to several others where participants rated video clips at the end of a workshop. A minimum of 50 participants is recommended for reliability studies (Cicchetti, 2001), but the sample in other similar studies was similar to or smaller than in this study (Lyk-Jensen et al., 2016; Rutherford et al., 2015; Spanager et al., 2013). Despite having slightly less than 50 participants, the sample included all those involved in clinical supervision of the SNAs on which the instrument is to be tested in a follow-up study. Thus, the participants were representative of the group for which the instrument is intended.

It is recommended that training courses for NTS last two full days (Hull et al., 2013; Klampfer et al., 2001). Attempts were made to increase both participation and the length of the workshop, but conflicting interests made following recommendations challenging. Other studies also found this impossible to implement due to staffing requirements, cost implications and effectivity in the operating department (Higham et al., 2019; Lyk-Jensen et al., 2016; Rutherford et al., 2015; Spanager et al., 2013). Another possible limitation was the course not being held by a multidisciplinary team of clinicians and psychologists/human factors experts (Flin & Maran, 2015; Hull et al., 2013), as in some other studies. However, the main researcher who held the course has both clinical expertise and considerable experience with teaching NTS.

It is generally considered to be easier to rate NTS in video clips than in clinical practice (Flin & Maran, 2015). Since the participants were not given any opportunity to rewind the clip, the scripted situations were over as quickly as real life situations. Other studies have used a larger number than six video clips, thereby providing a greater variety of situations. The video clips in this study were relatively long therefore six was deemed a sufficient number. However, increasing the number or length of the video clips can lead to rater fatigue and affect the results. Rater fatigue could have been lessened by using six shorter video clips, allowing more time for rating or spreading rating sessions over two days. However, such changes would either negatively affect the quality of the data or prove difficult to implement owing to organisational constraints.

Although the expert panel did not include a psychologist or expert in human factors, all the members had wide clinical experience as well as experience in teaching, training or assessing NTS. Knowledge of and experience in the field, inclination and time to participate as well as having effective communication skills, are regarded as the most important qualifications for members of an expert panel (Keeney et al., 2011). Since a classical Delphi approach was not followed to achieve consensus, panel members may have been influenced by each other during the face-to-face discussions. It is also possible that four is rather a small number for an expert panel, though there are no fixed guidelines for size and composition.

The evaluation questionnaire was translated to Norwegian, but a back translation was not deemed necessary since its purpose and content were very simple and the main researcher bilingual.

CONCLUSIONS

This study estimated a high level of reliability and dependability when rating NTS in video clips with the structured behavioural assessment instrument, NANTS-no. The findings suggest that there is a good foundation for further testing NANTS-no in clinical settings as a means of providing feedback and making structured assessments of student nurse anaesthetists. The findings also suggest that it may be useful for encouraging critical reflection in SNAs. Further research is needed to explore its use in clinical supervision.

Relevance to nursing practice and education

This study addressed the need for a formalized training and reliable instruments for use in clinical supervision of SNAs. This research suggests that NANTS-no has high reliability, dependability and acceptability in the study setting and may be useful as a structured assessment instrument for clinical supervision of student nurse anaesthetists in Norway.

References

Averlid, G. (2017). Norwegian nurse anesthetist perceptions of professional development and the influence of production pressure. *AANA J, 85*, 345-351.

Brennan, R. L. (2010). *Generalizability theory*. New York: Springer.

Cicchetti, D. V. (2001). The precision of reliability and validity estimates re-visited: distinguishing between clinical and statistical significance of sample size requirements. *J Clin Exp Neuropsychol, 23*(5), 695-700. doi:10.1076/jcen.23.5.695.1249

Collins, S., & Callahan, M. F. (2014). A call for change: clinical evaluation of student registered nurse anesthetists. *AANA J, 82*(1), 65.

Fletcher, G., Flin, R., McGeorge, P., Glavin, R., Maran, N., & Patey, R. (2003). Anaesthetists' Non-Technical Skills (ANTS): evaluation of a behavioural marker system. *Br J Anaesth, 90*(5), 580-588.

Flin, R., & Maran, N. (2015). Basic concepts for crew resource management and non-technical skills. *Best Pract Res Clin Anaesthesiol, 29*(1), 27-39. doi:10.1016/j.bpa.2015.02.002

Flin, R., & Mitchell, L. (2009). Safer surgery: analysing behaviour in the operating theatre. Farnham, England: Ashgate.

Flin, R., O'Connor, P., & Crichton, M. (2008). *Safety at the sharp end: a guide to non-technical skills*. Aldershot: Ashgate.

Flin, R., & Patey, R. (2011). Non-technical skills for anaesthetists: developing and applying ANTS. *Best Pract Res Clin Anaesthesiol, 25*(2), 215-227. doi:10.1016/j.bpa.2011.02.005

Flin, R., Patey, R., Glavin, R., & Maran, N. (2010). Anaesthetists' non-technical skills. *Br J Anaesth, 105*(1), 38-44. doi:10.1093/bja/aeq134

Flynn, F. M., Sandaker, K., & Ballangrud, R. (2017). Aiming for excellence – A simulation-based study on adapting and testing an instrument for developing non-technical skills in Norwegian student nurse anaesthetists. *Nurse Education in Practice, 22*, 37-46. doi:http://dx.doi.org/10.1016/j.nepr.2016.11.008

Fowler, J., & Cutcliffe, J. R. (2011). Clinical supervision: origins, overviews and rudiments. In J. R. Cutcliffe, Hyrkäs, K. & Fowler, J. (Ed.), *Routledge Handbook of Clinical Supervision*. Abingdon, Oxon: Routledge.

Gaba, D. M., Howard, S. K., & Small, S. D. (1995). Situation awareness in anesthesiology. *Hum Factors, 37*(1), 20-31.

Glavin, R. J. (2009). Excellence in anesthesiology: the role of nontechnical skills (Vol. 110, pp. 201).

Graham, J., Hocking, G., & Giles, E. (2010). Anaesthesia non-technical skills: Can anaesthetists be trained to reliably use this behavioural marker system in 1 day? *Br J Anaesth, 104*(4), 440-445. doi:10.1093/bja/aeq032

Haugen, A. S., Waehle, H. V., Almeland, S. K., Harthug, S., Sevdalis, N., Eide, G. E., . . . Softeland, E. (2017). Causal Analysis of World Health Organization's Surgical Safety Checklist Implementation Quality and Impact on Care Processes and Patient Outcomes: Secondary Analysis From a Large Stepped Wedge Cluster Randomized Controlled Trial in Norway. *Ann Surg*. doi:10.1097/sla.0000000000002584

Herion, C., Egger, L., Greif, R., & Violato, C. (2019). Validating international CanMEDS-based standards defining education and safe practice of nurse anesthetists. *Int Nurs Rev, 66*(3), 404-415. doi:10.1111/inr.12503

Higham, H., Greig, P. R., Rutherford, J., Vincent, L., Young, D., & Vincent, C. (2019). Observer-based tools for non-technical skills assessment in simulated and real clinical environments in healthcare: a systematic review. *BMJ Quality & Safety, 28*(8), 672-686. doi:10.1136/bmjqs-2018-008565

Hull, L., Arora, S., Symons, N. R., Jalil, R., Darzi, A., Vincent, C., & Sevdalis, N. (2013). Training faculty in nontechnical skill assessment: national guidelines on program requirements. *Ann Surg, 258*(2), 370-375. doi:10.1097/SLA.0b013e318279560b

IFNA. (2016). Standards of Education, Practice and Monitoring. Retrieved from http://ifna.site/download/ifna-standards-of-education-practice-and-monitoring-2016/
Jepsen, R. M., Dieckmann, P., Spanager, L., Lyk-Jensen, H. T., Konge, L., Ringsted, C., & Ostergaard, D. (2016). Evaluating structured assessment of anaesthesiologists' non-technical skills. *Acta Anaesthesiol Scand, 60*(6), 756-766. doi:10.1111/aas.12709
Jepsen, R. M., Østergaard, D., & Dieckmann, P. (2015). Development of instruments for assessment of individuals' and teams' non-technical skills in healthcare: a critical review. *Cognition, Technology & Work, 17*(1), 63-77. doi:10.1007/s10111-014-0306-y
Jha, A. K., Prasopa-Plaizier, N., Larizgoitia, I., & Bates, D. W. (2010). Patient safety research: an overview of the global evidence. *Qual Saf Health Care, 19*(1), 42-47. doi:10.1136/qshc.2008.029165
Jokelainen, M., Turunen, H., Tossavainen, K., Jamookeeah, D., & Coco, K. (2011). A systematic review of mentoring nursing students in clinical placements. *J Clin Nurs, 20*(19-20), 2854-2867. doi:10.1111/j.1365-2702.2010.03571.x
Jølstad, A. L., Røsnæs, E. R., Lyberg, A., & Severinsson, E. (2017). Clinical Supervision and Non-Technical Professional Development Skills in the Context of Patient Safety—The Views of Nurse Specialist Students. *Open Journal of Nursing, 7*, 254-267.
Keeney, S., Hasson, F., & McKenna, H. P. (2011). *The Delphi technique in nursing and health research*
Kennerly, D. A., Kudyakov, R., da Graca, B., Saldana, M., Compton, J., Nicewander, D., & Gilder, R. (2014). Characterization of adverse events detected in a large health care delivery system using an enhanced global trigger tool over a five-year interval. *Health Serv Res, 49*(5), 1407-1425. doi:10.1111/1475-6773.12163
Klampfer, B., Flin, R., Helmreich, R., Häusler, R., Sexton, B., Fletcher, G., . . . Amacher, A. (2001). *Enhancing Performance in High Risk Environments: Recommendations for the use of behavioural markers*. Retrieved from
Larsson, J., & Holmstrom, I. K. (2013). How excellent anaesthetists perform in the operating theatre: a qualitative study on non-technical skills. *Br J Anaesth, 110*(1), 115-121. doi:10.1093/bja/aes359
Lyk-Jensen, H. T., Dieckmann, P., Konge, L., Jepsen, R. M., Spanager, L., & Ostergaard, D. (2016). Using a Structured Assessment Tool to Evaluate Nontechnical Skills of Nurse Anesthetists. *AANA J, 84*(2), 122-127.
Lyk-Jensen, H. T., Jepsen, R. M., Spanager, L., Dieckmann, P., & Ostergaard, D. (2014). Assessing Nurse Anaesthetists' Non-Technical Skills in the operating room. *Acta Anaesthesiol Scand, 58*(7), 794-801. doi:10.1111/aas.12315
Lyth, G. M. (2000). Clinical supervision: a concept analysis. *J Adv Nurs, 31*(3), 722-729. doi:10.1046/j.1365-2648.2000.01329.x
Mitchell, L., Flin, R., Yule, S., Mitchell, J., Coutts, K., & Youngson, G. (2013). Development of a behavioural marker system for scrub practitioners' non-technical skills (SPLINTS system). *Journal of Evaluation in Clinical Practice, 19*(2), 317-323. doi:10.1111/j.1365-2753.2012.01825.x
Mushquash, C., & O'Connor, B. P. (2006). SPSS and SAS programs for generalizability theory analyses. *Behavior Research Methods, 38*(3), 542-547. doi:10.3758/bf03192810
Nilsson, U., & Jaensson, M. (2016). Anesthetic Nursing: Keep in Touch, Watch Over, and Be One Step Ahead. *J Perianesth Nurs, 31*(6), 550-551. doi:10.1016/j.jopan.2016.09.005
O'Donnell, D., Cook, N., & Black, P. (2016). Person-centred nursing education. In B. McCormack & T. McCance (Eds.), *Person-Centred Practice in Nursing and Health Care: Theory and Practice* (pp. 99-117). Oxford: John Wiley & sons.
Portney, L. G., & Watkins, M. P. (2009). *Foundations of Clinical Research: Applications to Practice*. Upper Saddle River: Pearson Prentice Hall.
Proctor, B. (1991). Supervision: A Co-Operative Exercise in Accountability. In M. a. P. Marken, M., (Ed.), *Enabling and Ensuring: Supervision in Practice* (pp. 21-23). Leicester: National Bureau and Council for Education and Training in Youth and Community Work.

Ringvold, E. M., Bekkevold, M., Bruun, A. G., Borke, W. B., Finjarn, T. J., Haugen, A. S., . . . Ulvik, A. (2018). Norwegian standard for the safe practice of anaesthesia. *Acta Anaesthesiol Scand, 62*(3), 411-417. doi:10.1111/aas.13066

Rutherford, J. S., Flin, R., Irwin, A., & McFadyen, A. K. (2015). Evaluation of the prototype Anaesthetic Non-technical Skills for Anaesthetic Practitioners (ANTS-AP) system: a behavioural rating system to assess the non-technical skills used by staff assisting the anaesthetist. *Anaesthesia, 70*(8), 907-914. doi:10.1111/anae.13127

Rutherford, J. S., Flin, R., & Mitchell, L. (2012). Non-technical skills of anaesthetic assistants in the perioperative period: a literature review. *Br J Anaesth, 109*(1), 27-31. doi:10.1093/bja/aes125

Rylance, R., Barrett, J., Sixsmith, P., & Ward, D. (2017). Student nurse mentoring: an evaluative study of the mentor's perspective. *Br J Nurs, 26*(7), 405-409. doi:10.12968/bjon.2017.26.7.405

Sevdalis, N., Hull, L., & Birnbach, D. J. (2012). Improving patient safety in the operating theatre and perioperative care: obstacles, interventions, and priorities for accelerating progress. *Br J Anaesth, 109*(suppl 1), i3-i16. doi:10.1093/bja/aes391

Smith, A. F., Glavin, R., & Greaves, J. D. (2011). Defining excellence in anaesthesia: the role of personal qualities and practice environment. *Br J Anaesth, 106*(1), 38-43. doi:10.1093/bja/aeq308

Spanager, L., Beier-Holgersen, R., Dieckmann, P., Konge, L., Rosenberg, J., & Oestergaard, D. (2013). Reliable assessment of general surgeons' non-technical skills based on video-recordings of patient simulated scenarios. *Am J Surg, 206*(5), 810-817. doi:10.1016/j.amjsurg.2013.04.002

Streiner, D. L., Norman, G. R., & Cairney, J. (2015). *Health measurement scales: a practical guide to their development and use* (5th ed. ed.). Oxford: Oxford University Press.

Tan, S. B., Pena, G., Altree, M., & Maddern, G. J. (2014). Multidisciplinary team simulation for the operating theatre: a review of the literature. *ANZ J Surg, 84*(7-8), 515-522. doi:10.1111/ans.12478

The Norwegian Association of Higher Education Institutions. (2016). *Kvalitet i praksisstudiene i helse- og sosialfaglig høyere utdanning* Retrieved from http://www.uhr.no/documents/praksisprosjektet_sluttrapport_ver2.pdf

Wisborg, T., & Manser, T. (2014). Assessment of non-technical skills in the operating room--one assessment tool per specialty? *Acta Anaesthesiol Scand, 58*(7), 773-774. doi:10.1111/aas.12352

Wong, A. (2012). Review article: teaching, learning, and the pursuit of excellence in anesthesia education. *Can J Anaesth, 59*(2), 171-181. doi:10.1007/s12630-011-9636-x

World Medical Association, W. (2013). Declaration of Helsinki. Retrieved from https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/

Yule, S., Flin, R., Paterson-Brown, S., Maran, N., & Rowley, D. (2006). Development of a rating system for surgeons' non-technical skills. *Med Educ, 40*(11), 1098-1104. doi:10.1111/j.1365-2929.2006.02610.x