ORIGINAL ARTICLE

Crew Emergency Teamwork Assessment Measure (CETAM): A Simulation-Based First Aid Study on Adapting and Validating a Tool for the Assessment of Nontechnical Skills in Airline Crew

Muhamad Nur Fariduddin[1] , Wee Lei Hum[2] , Lilia Halim[3], Mohd Johar Jaafar[4]

[1] Faculty of Education, Level 5 & 7, FSK Building 1,5, Universiti Teknologi MARA (UiTM), 42300, Bandar Puncak Alam, Selangor Darul Ehsan, Malaysia
[2] Faculty of Health Sciences, Universiti Kebangsaan Malaysia (UKM), Jalan Raja Muda Abdul Aziz, 50300 Kuala Lumpur, Malaysia
[3] Faculty of Education, Universiti Kebangsaan Malaysia (UKM), 43600 Bangi, Selangor Darul Ehsan, Malaysia
[4] Faculty of Medicine, Pusat Perubatan Universiti Kebangsaan Malaysia (UKM), Jalan Yaacob Latif, Bandar Tun Razak, 56000 Cheras, Wilayah Persekutuan Kuala Lumpur, Malaysia

ABSTRACT

Introduction: Diverse tools dedicated to the assessment of various NTS components during cardiopulmonary resuscitation have been published for medical professionals, nevertheless there are no precise measures for the assessment of emergency resuscitation relevant in the context of non-medical community. The aim of the study is to combine and develop a valid and reliable nontechnical skills (NTS) tool for basic life support emergency resuscitation training via simulation learning for flight attendant. **Methods:** (1) Selection and combination of relevant items into a draft instrument with a UKM specialist team. (2) Obtaining the expert's review from various specialty for content validity. (3) Instrument testing through pilot study on five recorded simulated case scenarios involving four different elements of nontechnical skills (teamwork, communication, decision making & situation awareness) which were demonstrated in a medical emergency (acute myocardial infarction) (4) Obtaining the internal consistency & inter-rater reliability of the instrument. **Results:** Through expert review, selected items had been found to have an excellent total content validity index of 1.00. A single mean of both raters yielded a good internal consistency of 0.77. Strong correlations of both ratings from raters in each video ranging from 0.69 to 0.93 ($p < .05$). There was a good inter-rater reliability (ICC 0.68) and a good agreement (Kappa 0.62 – 0.81) among raters. **Conclusion:** The CETAM was found to be a reliable and valid instrument and should be beneficial for the measurement of NTS, nevertheless further evaluation is required to fully determined its properties before reaching definitive conclusion.

Keywords: Nontechnical skills, Teamwork, Assessment, Flight attendant, Basic life support

Corresponding Author:
Wee Lei Hum, PhD
Tel: +603 92897902
Email: weeleihum@ukm.edu.my

INTRODUCTION

Nontechnical skills (NTS) are critical in preventing errors in healthcare (1-3). NTS covers good teamwork, communication, situation awareness and decision making (4). Errors occur in high stakes healthcare environments, such as the operating room or the trauma room, often owing the lack of nontechnical skills, therefore assuring that medical professionals possess adequate nontechnical skills is an essential part of medical education. There is a developing interest in the improvement of NTS and teamwork in medical practice as they are thought to increase patients' safety and care efficiency (5). For the past several years, realistic simulation methodology is presented as a methodological teaching resource that allows the individual to enter into a context that simulates. The simulation experience promotes critical thinking and acquisition of NTS which makes an advantageous strategy than traditional medical education (6, 7). Thus, many institutions have adopted training through simulation which allows practice and skill development where the effect of errors is minimal, moreover with the protected environment of the simulation, the student has the opportunity to learn by doing, making mistakes and learning from them (8).

In addition to improving nontechnical skills, there is a need to document the skill levels to ascertain whether they meet the standards and to evaluate the effectiveness of training. To foster such improvement in nontechnical skills, several assessments of nontechnical skills have been developed (1, 9). However, limited evidence

exists on the impact of healthcare professional's NTS development on patients' outcome. The main barrier for evaluating NTS impact on patients' safety is the absence of precise assessment techniques during simulation (5). In the aviation industry, flight attendant manages critical illness such as cardiac arrest and injury on a rare or unplanned basis in regular flight. Furthermore, nearly of all medical emergency cases had been handled alone by flight attendant, because the first individual at the scene of a medical emergency is often them instead of a medical practitioner (10, 11). With limited assistance within doubtlessly harsh condition prior to a definitive care, results in working in a setting which requires decisions to be made primarily based on the patient's chance of survival. For that reason, the importance of NTS within these settings is evident which advocate the necessities of high level of acquisition on nontechnical skills. Cardiopulmonary resuscitation (CPR) is the most frequently encountered medical situation over the last several years (12). For this reason, emphasis on the training has started to incorporate simulated clinical scenario to evaluate the NTS and team performance among flight attendants (13). In this context, diverse tools dedicated to the evaluation of various NTS components during cardiopulmonary resuscitation were published over the past decades (14, 15) but, it is crucial to note that the point of interest might not be relevant to the contexts of non-medical community, which include the aviation industry. To this point, no such behavior rating systems were developed and applied within the aviation industry. The use of such system allows reliable and effective observational evaluation of personnel engagement in passenger safety issues and measurement and benchmarking of behaviors associated with the implementation of NTS.

We intend to develop an easy to use tool to measure NTS and their contribution to universal team performance via the simulation conducted. Our aim was to ensure this tool can be use in the aviation industries to assist the training of BLS amongst flight attendants as a way to promote the acquisition of NTS amongst professionals in the non-medical community. As such, our aim was to evaluate the reliability and validity of the resulting crew Emergency Teamwork Assessment Measure (CETAM) which involved developmental stages: (i) combination and selection of items, and (ii) establishing the validity and reliability of the instruments based on the pilot testing of simulated emergency resuscitation by flight attendants.

METHODS

This project was developed by a team of health educationist from Universiti Kebangsaan Malaysia (UKM) together with a collaboration from one of the Airline Company in South East Asia. The instrument was developed in two stages: to begin with, an in depth review of the selected instrument particularly for teamwork among healthcare professionals were thoroughly screened by the panel of experts for the criteria selection of suitable items in order to develop a draft instrument which consist of NTS elements for emergency resuscitation teamwork. Next, the drafted instrument was then further review by the panel of experts for content validity followed by instrument testing on five recorded simulated emergency resuscitation (n = 5) for internal consistency and inter-rater reliability.

Crew Emergency Teamwork Assessment Measure (CETAM) Development

The Crew Emergency Teamwork Assessment Measure (CETAM) was developed based on the combination of elements from both validated instruments in assessing NTS in teamwork which is Clinical Teamwork Scale (CTS) (16) & Team Emergency Assessment Measure (TEAM) (15). According to the team training in emergency resuscitation among flight attendants in Asia context (27), the NTS skills emphasize has been given on four main elements which are effective teamwork among team members in delegating works and information pertaining to the emergency cases encountered, communication between crew and flight deck crew members in exchanging information and details, situational awareness and decision making. Thus, these four main elements were selected from both instruments and further combined into a single draft instruments. All of the items written with the descriptive definition as well as the rating scale was adopted. These instruments were reviewed by three different panel of experts in order to identify the applicable categories, suitability of the elements and items. This exercise led to a list of eleven items grouped under four different elements: teamwork (six items), communication (three items), decision making (one item) & situational awareness (one items) all rated on a scale of 0 to 10, and the scale values are further anchored by the qualitative descriptors of 'unacceptable' (0), 'poor' (1-3), 'average' (4-6), 'good' (7-9) and 'perfect' (10). We decided to remove the other elements from both instruments as these elements are not aligned with the current teaching and practice of NTS in the aviation academy related to the emergency resuscitation (Table I).

The next step design required expert content to review and rate the relevance of each items in the instruments based on a 4 point Likert scale.

Content Validity

The face and content validity of the CETAM had been assessed with a panel of three experts (EM, AE & ME) associated with the context of simulation learning amongst healthcare professionals with years of experience in the related discipline (Table II). Each of the selected experts was requested to independently rate the relevance of the eleven CETAM items using a 4-point scale (1 = not relevant, 2 = somewhat relevant, 3 = quite relevant, 4 = highly relevant). Level of agreement

Table I: Crew Emergency Teamwork Assessment Measure *(CETAM)* components and descriptive anchor

Elements	Component	Description
Teamwork	The team communicated effectively	Verbal, non-verbal & written forms of communication
	The team work together to complete task in a timely manner	-
	The team acted with composure & control	Applicable emotions/ conflict management issues
	The team morale was positive	Appropriate support, confidence, spirit, optimism, determination
	The team monitored & reassessed the situation	-
	The team anticipated potential actions	Preparation for defibrillator, airway equipment
Communication	Orient new member (SBAR)	As each new team member joined the scenario, they were oriented to the patient situation through a systematic communication using the format: S (Situation) B (Background) A (Assessment) R (Response)
	Directed communication	Team members assign requests (including orders) either verbally or visually to a specific person
	Closed loop communication	Team members acknowledge request and report back to the person issuing an order or requesting a specific action when the task is complete
Situational Awareness	Resource allocation	Team members vigilantly survey surroundings to be aware of all human and technological resources available and how to access them quickly
Decision Making	Prioritize	Clear, proper identification and ranking of items, actions, and/or issues pertinent to the management of the clinical situation

among experts was determined through calculating the Content Validity Index (CVI) for each individual item and the combined rating for the eleven items based on the proportion of experts with a rating of 3 and 4 (17).

With five or fewer experts, the value of validity index for each item (I-CVI) not lower than 0.78 can be considered a reasonable presentation of the universe feasible score (18) and the total content of the overall scale (S-CVI) with the value of .80 or higher is acceptable (19 – 21). Based on table II below, all items had a CVI of 1.00 (greater than 0.78) with a total item content of 1.00 (greater than 0.80) with the subsequent retention of all eleven objects.

CETAM Validity & Reliability Through Pilot Testing

The CETAM was pilot tested on flight attendants who attended a refresher training in their respective airline academy through a half-day session of first aid training. The group ranging from 10 to 15 members, however only five (n =5) members were selected to be part of the simulated resuscitation training. The selected team led one simulation based scenario specifically designed with the integration of knowledge, technical skills and most importantly, the elements of NTS. Prior to the event, all members consented to the study and briefed accordingly on the intention of the simulation training on the criteria and elements expected to be observed. The scenarios conducted are completely recorded for the purpose of video evaluation with the aid of the selected raters using CETAM. Two raters (anesthesiologist [AE], medical educationist [ME]) independently viewed the unidentified videos in no particular order and were asked to score the elements of NTS using the CETAM. In regard to the development of this tool, we do not emphasize the raters training session considering the fact that our primary aim was to develop a tool that would be utilized by the professionals in the particular area (aviation) for their personal assessment in normal practice. Aside from it, the professionals in aviation industry were trained in the principles of CRM (crew resource management) which is sufficient for knowledge and skills for human errors and team training (Fig 1).

Table II: Content Validity Index for *CETAM*

Item	Expert 1 (EM)	Expert 2 (AE)	Expert 3 (ME)	No of agreement	I-CVI
Teamwork (Component 1)	√	√	√	3	1.00
Teamwork (Component 2)	√	√	√	3	1.00
Teamwork (Component 3)	√	√	√	3	1.00
Teamwork (Component 4)	√	√	√	3	1.00
Teamwork (Component 5)	√	√	√	3	1.00
Teamwork (Component 6)	√	√	√	3	1.00
Communication (Comp 1)	√	√	√	3	1.00
Communication (Comp 2)	√	√	√	3	1.00
Communication (Comp 3)	√	√	√	3	1.00
Situational Awareness	√	√	√	3	1.00
Decision Making	√	√	√	3	1.00

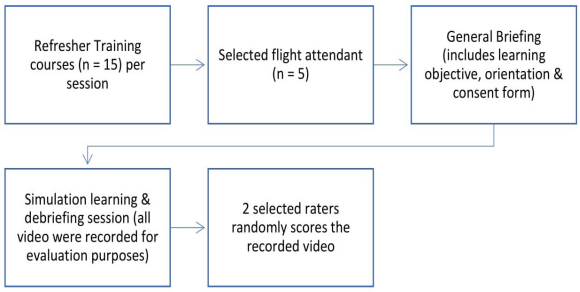

Figure 1: The summary of a simulation based program

Ethical Considerations

Ethics approval was obtained from the Faculty of Health Sciences, Universiti Kebangsaan Malaysia (UKM) Research Committee (NN-2017-105) on Ethical Research in humans. Following approval, flight attendants were given both oral and written information about the study and invited to participate by the author with the help of the academy administration team. In accordance with the ethical guidelines, the principles of informed consent, voluntary participation and the right to withdraw without penalty were cautiously explained to the participants, in addition to requirements regarding confidentiality, data anonymity and secure handling of video files. The written consent was obtained from all 30 flight attendants.

Data Analysis

Data analyses were carried out using IBM SPSS version 22.0 & MedCalc version 17.2. Most analysis were carried out by using IBMM SPSS except for the Weighted Kappa analysis which were carried out by using MedCalc. Content Validity was determined via calculating the Content Validity Index (CVI). Reliability of the CETAM was tested via Cronbach Alpha Coefficient for internal consistency, in addition to Intraclass Correlation Coefficient and Weighted Kappa technique for inter-rater reliability.

RESULTS

Reliability of items

To establish the internal consistency/reliability of the eleven individual CETAM items, a Cronbach Alpha coefficient (Table III) was calculated based on ratings of the five recorded-simulated resuscitation teamwork scenarios by both raters (AE & ME). Ideally, the Cronbach Alpha coefficient of a scale should be above 0.70 (22). This analysis on a single mean by both raters yielded an alpha coefficient of 0.77.

Inter-rater reliability

All of the five recorded-simulated resuscitation teamwork scenarios was used to determine the inter-rater reliability by examining the level of agreement between both raters on the scale scores of CETAM. A

Table III: Internal consistency of *CETAM* with Cronbach Alpha coefficient

Recorded video	Mean	SD	Cronbach's Alpha
1	4.28	.261	
2	4.28	.261	
3	4.09	.202	.773
4	3.90	.301	
5	3.90	.375	

Spearman Correlation Coefficient was used to determine the correlation between raters scores across the five different scenarios (Table IV). The overall item score correlations between raters for all five different scenarios were moderate to strong ranging between 0.69 to 0.93. The mean intra-class correlation coefficient of inter-rater reliability (Table V) was also good at 0.69 (23). The degree of agreement was also determined by using Weighted Kappa method whereby this method was more suitable to assess the rate of agreement for ordinal scales form of instruments (Table VI). As such, the values between 0.6 to 0.8 indicate good substantial agreement, and levels exceeding 0.8 are considered excellent (24). The overall agreement between the two raters (AE & ME) ranging from 0.62 to 0.81 which is good inter-observer agreement. When there was no agreement on each scale, there was only one-point difference in the ratings, for instance an allocation of '4' by one rater and '5' by another.

Table IV: Score correlations between pairs of raters across scenarios

Recorded Video	Correlation Coefficient *(rho)*	*P*
1	0.69	*< 0.05*
2	0.75	*< 0.05*
3	0.84	*< 0.05*
4	0.93	*< 0.05*
5	0.81	*< 0.05*

Table V: The Intraclass Correlation Coefficient of interrater reliability

	Intraclass Correlation[a]	95% Confidence Interval
Single measures [b]	0.180	0.044 to 0.477
Average measures [c]	0.687	0.314 to 0.901

[a] The degree of absolute agreement among measurements.
[b] Estimates the reliability of single ratings.
[c] Estimates the reliability of averages of k ratings.

Table VI: The degree of agreement between pairs of raters across scenarios

Recorded Video	Weighted Kappa	Interpretation
1	0.62	Good
2	0.65	Good
3	0.76	Good
4	0.81	Excellent
5	0.65	Good

DISCUSSION

With the rising medical emergencies reported onboard, it's far important to emphasize simulation training as part of the global suggestions to ensure patient safety and reducing preventable safety events. It's important to have tools which could evaluate the effectiveness of nontechnical skills training via simulation interventions, nevertheless the current available nontechnical skills measurement tools did not meet the requirement for aviation training since it was particularly tailored to the curriculum or the applications in which unit or medical institution workforce desire to evaluate their nontechnical skills in the course of emergencies or every day clinical care for both during simulation or daily exercise (1).

The Crew Emergency Teamwork Assessment Measure (CETAM) tools presented in this article were originally adopted and modified with a combination from two different tools, thus we tend to adhere to the original intention of the tools in which the tool designed as a universal measure for nontechnical elements which could be incorporated into different curriculum (15, 16). The value of Cronbach Alpha coefficient reported previously was 0.89 with an ICC value of 0.53, content validity index (CVI) of 0.96 & inter-rater agreement (Kappa) of 0.78 (15,16). In this context, it is proven that the feasibility and usability of the original elements works best in the aviation industry with several modifications on the definition of each element related to the aviation settings. It is also applicable to the aviation institution (the airline academy) which represents the teaching environments. This tool is intended to provide an overall evaluation by a group of flight attendant working together in a medical emergency through simulation exercise rather than several other nontechnical skill tools documented in the literature, in which majority of these tools were designed to evaluate nontechnical skills of individuals or as a team in a different and specific clinical settings (1).

Considering the fact that CETAM was developed using quantitative and qualitative research strategies with the input from the expertise range from skilled clinician and academicians, the outcomes suggest that the contents of the instruments are legitimate as validated by means of a high total Content Validity Index (1.00). The instruments have a good internal consistency (0.77) and a good inter-rater reliability demonstrated by the moderate to strong correlations of both ratings from raters for each element in each video ranging from 0.69 to 0.93. Apart from it, the Intraclass Correlation Coefficients (ICC) were good (0.68) as suggested by (23) based on the scale of 0.50 – 0.74 to be 'good' & above 0.75 to be 'excellent'. The inter-rater agreements from both raters using Weighted Kappa statistic were also 'good' & 'excellent' ranging from 0.62 – 0.81 based on the scale 0.61 – 0.80 'good' & 0.81 – 1.00 'excellent' (24) and whilst there was no agreement, there was only a single point difference in the rating scale for every item.

In the process of developing and validating this tools, like any other researcher, is not without any limitations. Based on the statistical output above, we would consider this tools to be feasible especially in the 'real time' testing but with some concerns over the issues such as the expertise of the selected raters. Since this tools concentrate on the elements of nontechnical skills rather than the technical skills aspects, it is unclear whether the raters from different profession, exposure and clinical experiences provided more reliable ratings (25). Though both raters had a good exposure and experience in conducting simulation learning, nevertheless this tools were developed specifically for aviation industry whereby the participants for the simulation interventions were 'layman' in a non-medical community. On top of it, the setting for this simulation training represent the medical emergencies in an aircraft setting which we find it a barrier for the raters to have better understanding of the situation to provide better ratings. It is possible that the amount of experience in an occupation may influence the quality of the ratings which can be seen based on value of Kappa obtained above (26). Unfortunately, most studies failed to investigate whether experience was related to ratings quality but in this context, we suggest rater training as a vital element for obtaining accurate and reliable ratings.

The scenarios designed and implemented in this study should be designed to provide multiple opportunities to observe behavior of interest. Typically, a longer scenario contains more opportunities for observation, unless each of the elements were specifically incorporated on each segments of the scenario. In this context, a 6 – 10 minutes' duration of scenario with high element of interactivity was designed to ensure the nontechnical skills can be observed, unfortunately the setting of this scenario involved relocating the victim from the aircraft seats, immediate chest compression and utilization of AED automatically affects the participants to act on the technical skills rather than the nontechnical skills. It is difficult for the rater to assess the relevant elements distinctly in a short time frame with minimal behavior expressed in the scenes. As such, we estimated more specifically designed scenarios for the nontechnical skills elements together with more number of scenarios to be observed by the raters to increase the reliability of the ratings.

CONCLUSION

Following development and preliminary testing in a simulated environment, the CETAM has emerged as a valid, reliable and feasible nontechnical skills measurement tool for the assessment of basic life support training for flight attendant related to emergency resuscitation performance. The instruments will allow

team overall performance assessment and feedback via simulation training on flight attendants that's likely to improve their performances with the intention to have an impact on patient's safety onboard. To that end, we have now applied the tool for the assessment of nontechnical skills training in an emergency heart attack simulated case scenario in length of 10 minutes over 150 flight attendants. With that, we discovered it useful in assisting the aviation team to conduct debriefing and we're continuing to evaluate the tool to achieve further information before attaining definitive conclusions.

ACKNOWLEDGEMENTS

The authors would like to thank all the flight attendants of the respective airline company who participated in the study, Madam Rosnina Abdullah, the Manager of Safety & Emergency Department for her valuable assistance throughout the process, Madam Shuraiya & Madam Eswari, the registered nurse and first aid instructor for their support and help with carrying out the study and the respective experts Mr Balakrishan (EM), Dr Jalina Karim (ME) & Dr Noorjahan (AE) for content validation and ratings.

DECLARATION OF INTEREST

The authors whose names are listed immediately above certify that they have no affiliations with or involvement in any organization or entity with any financial or non-financial interest in the subject matter or materials discussed in this manuscript.

REFERENCES

1. Jepsen RMHG, Шstergaard D, Dieckmann P. Development of instruments for assessment of individuals' and teams' non-technical skills in healthcare: A critical review. Cogn Tech Work. 2015;17:63.
2. Odell M. Human factors and patient safety: changing roles in critical care. Aust Crit Care. 2015;24(4):215-7.
3. Reader, Tom W, Flin R, Cuthbertson, Brian H. Communication skills and error in the intensive care unit. Cur Op Crit C. 2007;13(6):732 – 736.
4. Gordon M, Darbyshire D, Baker P. Non-technical skills training to enhance patient safety: a systematic review. Med Educ. 2012;46(11):1042-1054
5. Grand JA, Pearce M, Rench TA, et al. Going DEEP: guidelines for building simulation-based team assessments. BMJ Qual Saf. 2013;22:436-448.
6. Issenberg BS, Megaghie WC, Petrusa ER, Gordon DL, Scalese RJ. Features and uses of high fidelity medical simulations that lead to effective learning: A BEME systematic review. Med T. 2005;27(1):10-28
7. McGaghie WC, Issenberg SB, Petrusa ER, Scalese RJ. A critical review of simulation based medical education research: 2003-2009. Med Educ. 2010;44:50-63
8. Varga CRR, Almeida VC, Germano CMR, Melo DG, Chacha SGF, Souto BGA, et al. Relato de experiencia: o uso de simulacoes no processo de ensino-aprendizagem em medicina. Rev Bras Educ Med. 2009;33(2):291-297.
9. Manser T. Teamwork and patient safety in dynamic domains of healthcare: A review of the literature. Acta Anaesthesiol Scand. 2009;53(2):143–151.
10. Dowdall N. Is there a doctor on the aircraft? Top 10 in-flight medical emergencies. BMJ. 2007;321:1336-7
11. Hung KK, Cocks RA, Poon WK, Chan EY, Rainer TH, Graham CA. Medical volunteers in commercial flight medical diversions. Aviat Space Environ Med 2013;84:491–7
12. Amit C, Shauna C. In flight medical emergencies. Wes J Em Med. 2013;14(5):499-504.
13. Aziz FA, Osman K. Integrating Simulation Innovatively: Evidence in Teaching First Aid for Cabin Crew. Int J Adv App Sci. 2015;3(6):209-215.
14. Walker S, Brett S, McKay A, Lambden S, Vincent C, Sevdalis N. Observational skill-based clinical assessment tool for resuscitation (OSCAR): development and validation. Res 2011;82(7):835–844.
15. Cooper S, Endacott R, Cant R. Measuring non-technical skills in medical emergency care: a review of assessment measures. Em Med. 2010;2:7-16
16. Guise JM, Deering SH, Kanki BG, Osterweil P, Li H, Mori M, Lowe NK. Validation of a tool to measure and promote clinical teamwork. Simul Healthc 2008;3(4):217–223
17. Polit D, Beck C. The content validity index: are you sure you know what's being reported? Critique and recommendations. Res Nurs Health. 2006;29:489–97.
18. Lynn MR. Determination and quantification of content validity. Nurs Res. 1986;35:382-285
19. Davis LL. Instrument review: getting the most from your panel of experts. App Nurs Res. 1992;5:194-197
20. Grant JS, Davis LT. Selection and use of content experts in instrument development. Res Nurs Heal. 1997;20:269-274
21. Polit DF, Beck CT. Nursing research: Principles and methods (7th Ed). Philadelphia: Lippincott, Williams, & Wilkins; 2004.
22. DeVellis R. Scale development: theory and applications. Thousand Oaks: Sage Publications; 2003.
23. Fleiss JL. Measuring nominal scale agreement among many raters. Psy Bull. 1986;76:378–82.
24. Altman DG. Practical Statistics for Medical Research. London: Chapman & Hall ;1991 p.404.
25. Robertson ER, Hadi M, Morgan LJ, et al. Oxford

NOTECHS II: a modified theatre team non-technical skills scoring system. PLoS One. 2014;9(3):e90320.
26. Mete I, Michael TB. Estimating the reliability of nontechnical skills in medical teams. J of Surg Educ. 2016;74(4):596-611.
27. Malaysia Airlines Berhad. First Aid: Safety & Emergency Procedures Manual. MAB: Malaysia Airlines Academy, 2016.