

# High fidelity simulation to assess task load index and performance: a prospective observational study

**Jérémy Favre-Félix**

Hospices Civils de Lyon

**Mikhail Dziadzko**

Hospices Civils de Lyon

**Christian Bauer**

Hospices Civils de Lyon

**Antoine Duclos**

Hospices Civils de Lyon

**Jean-Jacques Lehot**

Hospices Civils de Lyon

**Thomas Rimmelé**

Hospices Civils de Lyon

**Marc Lilot** (✉ [marclilot@hotmail.com](mailto:marclilot@hotmail.com))

Hospices Civils de Lyon

---

## Research Article

**Keywords:** Anaesthesia, Clinical performance, High-fidelity simulation, Task-load index

**Posted Date:** February 25th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-236458/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

## Background

The NASA Task Load Index (NASA-TLX) is a questionnaire widely used in aviation. This index might help for attesting the quality of a scenario in high-fidelity simulation (HFS) in healthcare.

The main purpose of this study was to observe whether the score of NASA-TLX for critical care simulated scenarios, designed for residents, was consistent with the values reported in the literature.

The second purpose was to describe relationships between NASA-TLX, performance and generated stress during HFS.

## Methods

All residents in anaesthesia and intensive care undergoing HFS were included. The primary endpoint was the task load generated by each scenario assessed by NASA-TLX. Based on literature, the NASA-TLX scores between 39 and 61 were considered as acceptable level. Stress level (Visual Analogue Scale for stress), specific technical and non-technical skills performances (Team Emergency Assessment Measure) were also assessed.

## Results

Fifty-three residents actively participated in one of ten different scenarios, between June to December 2017. Median NASA-TLX score of scenarios was 61 [48–65]. Five scenarios (50%) generated acceptable task load level. There was no association between NASA-TLX score and technical or Team Emergency Assessment Measure performance scores, but an association between NASA-TLX and the stress level ( $\rho = 4.7$ ,  $p = 0.001$ ) was observed.

## Conclusions

Simulation scenarios generate different task loads for residents; the task load was deemed acceptable for half of the scenarios. The NASA-TLX could be considered as a tool to assess the pedagogic adequacy of scenarios. Scenario and generated stress level, but not task load, can modify residents' performance during simulation. This should be considered when planning normative simulation.

## Introduction

High-fidelity simulation (HFS) is used as an effective teaching method to enhance the acquisition of required competences in anaesthesiology and intensive care [1, 2]. By taking part in HFS scenarios, residents are actively involved in relevant critical situations [3]. HFS increases stress, mobilizing a level of

mental and physical resources which may be expressed as the task load [4, 5]. The relationship between stress and efficient memorization was reported to look like an inverted U shape function [6]. Too little stress would not help optimal learning whereas too much stress would prevent the student from memorizing relevant information [7]. Therefore, the scenario should be assessed for its potential task load generation and associated increase in stress level. Similarly, one may hypothesize that the optimal task load would be associated with higher performance, and the performance may be reduced in situations when the task load is too low or too high.

Different behavioural and subjective metrics are used in operational environments (aviation, nuclear power plants, medicine) to assess the task load of operators [8–11]. The NASA Task Load Index (NASA-TLX), initially developed for flight training, is the most commonly applied tool to assess procedural workload in healthcare and has been specifically evaluated in the surgical field [12]. This tool evaluates perceived task load immediately after the performed task through mental, emotional, and physical dimensions.

We hypothesized that the task load in anaesthesia and intensive care residents generated by the scenarios was consistent with the expected normal values reported in the literature.

The main objective of this study was to explore the task load (with the NASA-TLX) of each scenario during HFS for anaesthesia and intensive care residents. Secondary objectives were to describe relationships between NASA-TLX, scenario performance, and associated stress level.

## Methods

### *Design*

This observational prospective cohort study was conducted at the university medical simulation centre of Lyon (Lyon1 University, France). The study obtained approval from the Hospices Civil de Lyon institutional ethics committee (27/06/2017) and has been pre-registered on clinicaltrials.gov (Protocol ID: NCT03175484). An informed written consent was obtained from all enrolled participants. This research has been carried out in accordance with The Code of

Ethics of the World Medical Association (Declaration of Helsinki).

### *Population and simulation setting*

This study involved all residents in anaesthesia and intensive care undergoing HFS sessions between June and December 2017. No exclusion criterion was applied. The scenarios simulated crisis situations occurring in the operating room, in the intensive care unit or during intra-hospital patient transport. Scenarios were developed by instructors based on national guidelines. HFS sessions followed the standard repetitive sequences of briefing, scenario and debriefing.

### *2.3 Experimental protocol*

The demographic data of participants at the beginning of the HFS session were collected (age, gender, previous HFS participation, post graduate year). We used the NASA-TLX questionnaire to assess the task load of each participant during each HFS scenario. NASA-TLX score (0 point: no task load, 100 points: maximal task load) included six dimensions (mental demand, physical demand, temporal demand, performance, effort, frustration). A detailed description and the questionnaires used are reported in Appendix 1. As suggested in the literature, the NASA-TLX score was considered as consistent with those values if situated between 25<sup>th</sup> and 75<sup>th</sup> percentile ( $39 < \text{NASA-TLX} < 61$ ) [13]. We considered those NASA-TLX values as “acceptable”.

We measured the individual quantitative stress level using a Visual Analogue Scale for stress translated on a 100 mm numeric scale for stress (0: no stress, 100: maximal stress) immediately after each scenario. We also measured both technical performance, using a specific technical skills scoring grid (previously described [14], 0: low performance, 100: maximal performance) and non-technical skills performance by the Team Emergency Assessment Measure scale (0: low performance, 44: maximal performance) [15,16]. Two investigators (C.B., M.L.) independently evaluated the resident performance using video recording. The timeline of the study is presented in Figure 1.

#### *2.4 Endpoints*

The primary endpoint was the assessment of individual NASA-TLX score for each scenario. Secondary endpoints were the stress level at the end of the scenario measured by Visual Analogue Scale for stress, the technical and the Team Emergency Assessment Measure performance score.

#### *2.5 Statistical analysis*

Continuous variables were described using median [25<sup>th</sup>-75<sup>th</sup> percentile] and were compared using the Wilcoxon or the Student t test as appropriate. The Kruskal-Wallis test was used for more than two group comparisons. A multivariate linear regression analysis was performed to explore the interaction between NASA-TLX, scenario, stress and technical performance. Correlation between technical performance and non-technical performance skills scores was evaluated using Spearman’s (rho) correlation index. All tests were two-tailed, and  $p < 0.05$  was considered statistically significant. Statistical analysis was performed on a *perprotocol* basis using MedCalc software version 9.6.4.0 (MedCalc, Mariakerke, Belgium).

## **Results**

A total of 53 residents (mean age 26 years old, standard deviation = 2; 21 (40%) females), coming through HFS and involved in one of ten different scenarios were included and analysed from June to December 2017. The median NASA-TLX score of the ten scenarios was 61 [48 - 65]. Five scenarios (50%) generated NASA-TLX between 25<sup>th</sup> and 75<sup>th</sup> percentile (39 - 61 points), and five scenarios generated higher task load level (> 61 points). The median NASA-TLX score of the 53 residents for all scenarios was 61.7 [46.7 - 67.1], 16 participants (30%) had a score between 25<sup>th</sup> and 75<sup>th</sup> percentile (39 - 61 pts), eight (15%) had a lower score, and 29 (55%) had a higher score.

The median Visual Analogue Scale for stress was 43.0 [24.5 - 66.0], the median technical skills performance was 44.0 [34.0 - 52.5], and median Team Emergency Assessment Measure score was 23.5 [20.0 - 28.0]. There was no age or gender difference observed in NASA-TLX, Visual Analogue Scale for stress and performance scores. Except for the Visual Analogue Scale for stress, NASA-TLX and performance scores were significantly different across performed scenarios (Table 1).

The technical skills performance was correlated with the non-technical skills performance ( $\rho = 0.6$ ;  $p < 0.001$ ). In a bivariate fit analysis, NASA-TLX score was not associated with technical skills performance and Team Emergency Assessment Measure scores, but was associated with stress level (4.7 points of increment for 1 point increment of NASA-TLX score,  $p = 0.001$ , Figure 2).

An elevated Visual Analogue Scale for stress predicted decreased technical skills performance (-15 points decrement for each 10 points Visual Analogue Scale for stress increment,  $p = 0.03$ ). No significant association between Visual Analogue Scale for stress and Team Emergency Assessment Measure score was observed.

In multivariate analysis, the Visual Analogue Scale for stress incurred by HFS (F ratio = 4.16,  $p = 0.048$ ) and the type of scenario (F ratio = 17.6,  $p < 0.0001$ ) predicted technical skills performance. For the Team Emergency Assessment Measure performance, only the type of scenario (F ratio = 7.99,  $p < 0.0001$ ) was predictive (Appendix 2).

## Discussion

The task load level was assessed with the NASA-TLX score in 53 anaesthesia and intensive care residents undergoing HFS training. Among the scenarios tested, a half generated an adapted task load, and another half generated high task load level. Such scenarios are of great educational value but should be used with caution and may be suited to more experienced residents.

To the best of our knowledge, no study reporting a relationship between NASA-TLX and HFS performance scores in healthcare has been published. Different scales are available to measure task load. Subjective Workload Assessment Technique uses three levels (low, medium and high) for each of three dimensions of time load, mental effort load and psychological stress. Overall Workload allows the subjects rating overall workload on a unidimensional scale from 0 to 100 points. Studies which compared all of these validated scales demonstrated the reliability and validity of the NASA-TLX in comparison to other workload measures [17], probably due to the six subscales which allows more precision in the task load assessment. *Hart et. al* showed that NASA-TLX was the most used subjective scale to assess task load [18]. Previous validation in medical field has been published providing information to help interpreting the obtained scores for each scenario [13]. However, to the best of our knowledge, no study validated objective methods to assess task load in medical field.

The use of task load assessment by NASA-TLX was first described in aviation. It was demonstrated that reducing the task load significantly improved performance in aircraft [19]. In healthcare, an increasing number of studies focus on the task load measurement, since overload has been identified as a significant cause of errors [20, 21]. Increased response time during simulated crisis situations has also been reported with task overload [22, 23]. Residents might experience high task load during scenarios and such this may harm the process of learning. By contrast, scenarios with too low task load might result in poor involvement. Task load analysis becomes important when the cognitive load theory is considered. This theory assumes that working memory capacity is limited. In some complex learning cases, reducing task load will help increase working memory capacity [24]. In the present study, no extreme NASA-TLX scores (> 77 points) were observed, a level reported as clear overload in the literature [13].

An association between NASA-TLX and stress level was observed, while no association was observed with the technical or the non-technical performance. Those results suggest that NASA-TLX could be used as a marker of scenarios' educational quality but not as a tool to assess residents' performance during HFS. Similar results were found in a surgical simulation study exploring NASA-TLX during laparoscopy [25], which might be explained by the fact that performance and task load are two separate dimensions of the scenario. HFS scenario might be perceived as a challenge and with a need to mobilize new working resources, such as providing leadership during critical situations. These tasks will affect the task load without systematically affecting performance skills. The performance scores obtained during HFS results from multifactorial and complex factors that affect the participants before and during the scenario. The performance during the scenario is not a pedagogical objective of the HFS, but rather a tool to achieve better further performance in real setting (helped with the information provided during the debriefing). Thus, we might assume that performance during HFS is not, in this study, a valuable independent marker for scenario's educational quality. Moreover, the perceived performance is only one component of the six scales of NASA-TLX questionnaire and this might explain in parts why performance was not associated with NASA-TLX. The performance NASA-TLX subscale reflects the subjective self-evaluation of residents that could be influenced by several psychometric or emotional factors (self-efficacy feeling, fear of negative evaluation, social anxiety, reaction of participants to the announcement of the end of the scenario) [5, 26–28]. This auto-appreciation does not fit with the performance scale relying on specific objectives rated by investigators.

The association between task load and stress was previously described in surgery residents. NASA-TLX was positively correlated with objective stress levels measured by sympathetic activity (heart rate, blood pressure) [29]. These changes may be explained by specific NASA-TLX subscales as frustration, temporal demand, which are feelings of being overpassed which could increase stress level.

This study has several limitations. The number of residents was 53 and could have precluded the inference of NASA-TLX and performance. Ten scenarios for three different post graduate years was included offering variability that might have influenced the results. Then, as Hart *et al.* reported previously [18], we noticed that the main limitation of the NASA-TLX is the interpretation of the score. Thresholds used in this study were given by the analysis of the vast amounts of data published in the medical field

[13]. This analysis, mainly in simulated endoscopic surgery and emergency room situations did not match the setting for the participants (alone or in a team), and the performance did not impact on any certification. All of these differences could have influenced the task load and it is very difficult to establish a universal and reliable threshold in the medical field. However, identify the relative task load for each scenario might help instructors to identify specific interest of scenario in regards to the task load provided. Further studies are needed to assess the potential of errors production associated with perceived scenario-specific task load, to further evaluate the hypothetical difficulty of each scenario. There is a need for precision on the optimal NASA-TLX threshold to define locally what is an acceptable, high or low workload, in order to use this tool effectively to enhance the pedagogical values of HFS. NASA-TLX might be used as a tool to identify scenarios with outliers or marginal scores in order to select, upgrade or adapt scenarios to the pedagogical objectives of HFS.

## Conclusion

To conclude, simulation scenarios generate different task loads in residents and NASA-TLX could be considered as an additional tool to help instructors to assess the pedagogic adequacy of HFS scenarios to learners. Scenario and generated stress level, but not task load, can modify residents' performance during simulation. This should be considered when planning normative simulation.

## Abbreviations

HFS

High-fidelity simulation

NASA-TLX

NASA Task Load Index

## Declarations

- Ethics approval and consent to participate

The study obtained approval from the Hospices Civil de Lyon institutional ethics committee (27/06/2017) and has been pre-registered on [clinicaltrials.gov](https://clinicaltrials.gov) (Protocol ID: NCT03175484). An informed written consent was obtained from all enrolled participants. This research has been carried out in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki).

- Consent for publication

Not applicable. The authors declare that no information compromising anonymity appears in the article.

- Availability of data and materials

The datasets collected, used and analysed during the current study are available from the corresponding author on reasonable request.

- Competing interests

The authors declare that they have no competing interests.

- Funding

This work was supported by institutional and/or departmental resources.

No financial disclosure to declare. No non-financial disclosure to declare.

- Authors' contributions

Jérémy Favre-Félix, Mikhail Dziadzko, Christian Bauer, Antoine Duclos, Jean-Jacques Lehot, Thomas Rimmelé and Marc Lilot realized the study design.

Jérémy Favre-Félix, Christian Bauer, and Marc Lilot collected data. Jérémy Favre-Félix, Mikhail Dziadzko, Antoine Duclos and Marc Lilot performed the data analysis and the writing up of the first draft of the paper.

Jérémy Favre-Félix, Mikhail Dziadzko, Jean-Jacques Lehot, Thomas Rimmelé and Marc Lilot prepared the Figure the Table and the Appendices.

Jérémy Favre-Félix, Mikhail Dziadzko, Christian Bauer, Antoine Duclos, Jean-Jacques Lehot, Thomas Rimmelé and Marc Lilot revised the manuscript for substantial intellectual contents of the final manuscript, approved the final manuscript and agreed to be accountable for all aspects of the work thereby ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

- Acknowledgments:

The authors thank the instructors of the CLESS for their help in conducting the study: B. Bui-Xuan (MD), A. Faure (MD), and T. Geffriaud (MD). They thank Mr. L. Denoyel (CLESS, Lyon, France) for his technical support during sessions, and finally, the anaesthesiology and intensive care residents from Lyon 1 University for their participation.

## References

[1] Lorello GR, Cook DA, Johnson RL, Brydges R. Simulation-based training in anaesthesiology: a systematic review and meta-analysis. *British Journal of Anaesthesia* 2014; 112 : 231 – 45. <https://doi.org/10.1093/bja/aet414>.

- [2] Cook DA, Hatala R, Brydges R, Zendejas B, Szostek JH, Wang AT, et al. Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. *JAMA* 2011; 306 : 978 – 88. <https://doi.org/10.1001/jama.2011.1234>.
- [3] Issenberg SB, McGaghie WC, Petrusa ER, Lee Gordon D, Scalese RJ. Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. *Med Teach* 2005; 27 : 10 – 28. <https://doi.org/10.1080/01421590500046924>.
- [4] Bauer C, Rimmelé T, Duclos A, Prieto N, Cejka J-C, Carry P-Y, et al. Anxiety and stress among anaesthesiology and critical care residents during high-fidelity simulation sessions. *Anaesth Crit Care Pain Med* 2016; 35 : 407 – 16. <https://doi.org/10.1016/j.accpm.2016.01.004>.
- [5] Nielsen B, Harder B. Causes of Student Anxiety during Simulation: What the Literature Says. *Clinical Simulation in Nursing* 2013; 9 : e507 – e512. <https://doi.org/10.1016/j.ecns.2013.03.003>.
- [6] Salehi B, Cordero MI, Sandi C. Learning under stress: the inverted-U-shape function revisited. *Learn Mem* 2010; 17 : 522 – 30. <https://doi.org/10.1101/lm.1914110>.
- [7] LeBlanc VR. The effects of acute stress on performance: implications for health professions education. *Acad Med* 2009; 84 : S25 - 33. <https://doi.org/10.1097/ACM.0b013e3181b37b8f>.
- [8] Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. *Advances in Psychology*, vol. 52, Elsevier; 1988, p. 139 – 83. [https://doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9).
- [9] Eitrhein MH, Fernandes A. The NASA Task Load Index for rating workload acceptability. 2016. <https://doi.org/10.13140/RG.2.2.22989.64484>.
- [10] Wilson MR, Poolton JM, Malhotra N, Ngo K, Bright E, Masters RSW. Development and validation of a surgical workload measure: the surgery task load index (SURG-TLX). *World J Surg* 2011; 35 : 1961 – 9. <https://doi.org/10.1007/s00268-011-1141-4>.
- [11] Porges SW. The Polyvagal Perspective. *Biol Psychol* 2007; 74 : 116 – 43. <https://doi.org/10.1016/j.biopsycho.2006.06.009>.
- [12] Dias RD, Ngo-Howard MC, Boskovski MT, Zenati MA, Yule SJ. Systematic review of measurement tools to assess surgeons' intraoperative cognitive workload. *Br J Surg* 2018; 105 : 491 – 501. <https://doi.org/10.1002/bjs.10795>.
- [13] Grier RA. How High is High? A Meta-Analysis of NASA-TLX Global Workload Scores. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 2015; 59 : 1727 – 31. <https://doi.org/10.1177/1541931215591373>.

- [14] Evain J-N, Perrot A, Vincent A, Cejka J-C, Bauer C, Duclos A, et al. Team planning discussion and clinical performance: a prospective, randomised, controlled simulation trial. *Anaesthesia* 2019; 74 : 488 – 96. <https://doi.org/10.1111/anae.14602>.
- [15] Cooper S, Cant R, Porter J, Sellick K, Somers G, Kinsman L, et al. Rating medical emergency teamwork performance: development of the Team Emergency Assessment Measure (TEAM). *Resuscitation* 2010; 81 : 446 – 52. <https://doi.org/10.1016/j.resuscitation.2009.11.027>.
- [16] Maignan M, Koch F-X, Chaix J, Phellouzat P, Binauld G, Collomb Muret R, et al. Team Emergency Assessment Measure (TEAM) for the assessment of non-technical skills during resuscitation: Validation of the French version. *Resuscitation* 2016; 101 : 115 – 20. <https://doi.org/10.1016/j.resuscitation.2015.11.024>.
- [17] Hill SG, Iavecchia HP, Byers JC, Bittner AC, Zaklade AL, Christ RE. Comparison of Four Subjective Workload Rating Scales. *Hum Factors* 1992; 34 : 429 – 39. <https://doi.org/10.1177/001872089203400405>.
- [18] Hart SG. Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 2006; 50 : 904 – 8. <https://doi.org/10.1177/154193120605000909>.
- [19] Gruenwald CM, Middendorf MS, Hoepf MR, Galster SM. Augmenting Human Performance in Remotely Piloted Aircraft. *Aerosp Med Hum Perform* 2018; 89 : 115 – 21. <https://doi.org/10.3357/AMHP.4894.2018>.
- [20] Weinger MB, Reddy SB, Slagle JM. Multiple measures of anesthesia workload during teaching and nonteaching cases. *Anesth Analg* 2004; 98 : 1419 – 25, table of contents.
- [21] Yurko YY, Scerbo MW, Prabhu AS, Acker CE, Stefanidis D. Higher mental workload is associated with poorer laparoscopic performance as measured by the NASA-TLX tool. *Simul Healthc* 2010; 5 : 267 – 71. <https://doi.org/10.1097/SIH.0b013e3181e3f329>.
- [22] Davis DHJ, Oliver M, Byrne AJ. A novel method of measuring the mental workload of anaesthetists during simulated practice. *Br J Anaesth* 2009; 103 : 665 – 9. <https://doi.org/10.1093/bja/aep268>.
- [23] Astin A, Nussbaum MA. Interactive Effects of Physical and Mental Workload on Subjective Workload Assessment. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 2002; 46 : 1100 – 4. <https://doi.org/10.1177/154193120204601320>.
- [24] van Merriënboer JJG, Sweller J. Cognitive load theory in health professional education: design principles and strategies. *Med Educ* 2010; 44 : 85 – 93. <https://doi.org/10.1111/j.1365-2923.2009.03498.x>.
- [25] Sakata S, Grove PM, Watson MO, Stevenson ARL. The impact of crosstalk on three-dimensional laparoscopic performance and workload. *Surg Endosc* 2017; 31 : 4044 – 50.

<https://doi.org/10.1007/s00464-017-5449-5>.

[26] Watson D, Friend R. Measurement of social-evaluative anxiety. *J Consult Clin Psychol* 1969; 33 : 448 – 57. <https://doi.org/10.1037/h0027806>.

[27] Dieckmann P, Krage R. Simulation and psychology: creating, recognizing and using learning opportunities. *Curr Opin Anaesthesiol* 2013; 26 : 714 – 20. <https://doi.org/10.1097/ACO.0000000000000018>.

[28] Ozer EM, Bandura A. Mechanisms governing empowerment effects: a self-efficacy analysis. *J Pers Soc Psychol* 1990; 58 : 472 – 86. <https://doi.org/10.1037//0022-3514.58.3.472>.

[29] Bakhsh A, Martin GFJ, Bicknell CD, Pettengell C, Riga C. An Evaluation of the Impact of High-Fidelity Endovascular Simulation on Surgeon Stress and Technical Performance. *J Surg Educ* 2019; 76 : 864 – 71. <https://doi.org/10.1016/j.jsurg.2018.10.015>.

## Table

Topic of the scenario	Summary of the briefing	PGY	Male/ Female (n/n)	NASA- TLX	Technical performance	Non- technical skills performance	S-VAS
Announcement of a surgery side error	In the PACU, you receive the patient's family to explain surgery side error	5 n=4	2/2	46 [38-64]	41 [36-53]	20 [19-28]	40 [13-68]
Cardiogenic shock	You admit in the ICU a young man with acute hypotension, tachycardia and moderate fever context	5 n=4	2/2	64 [63-75]	56 [47-62]	33 [21-38]	41 [36-62]
Severe postpartum hemorrhagic shock	You are called for a woman with severe bleeding after delivery	5 n=4	2/2	65 [61-70]	76 [47-62]	34 [29-37]	32 [22-61]
Cardiac arrest due to local anesthetics toxicity	You are in charge of a patient under regional anaesthesia for a fracture fixation	5 n=4	2/2	45 [40-70]	61 [55-67]	36 [34-37]	18 [14-64]
Acute neurological disorders due to local anesthetics toxicity	You are in charge of a patient under regional anaesthesia for a fracture fixation	2 n=4	4/0	64 [62-72]	31 [27-36]	20 [19-24]	67 [38-84]
Gas embolism	You are in charge of a patient with spinal surgery under general anaesthesia, when acute dyspnea occurs	2 n=4	4/0	53 [33-56]	39 [29-45]	21 [19-26]	58 [26-70]
Acute post reperfusion ventricular fibrillation	You are in charge of a patient for of acute lower limb ischemia surgery	2 n=5	3/2	62 [48-73]	36 [34-45]	28 [23-29]	47 [33-80]
Acute hypoxia due to selective intubation	You are called to transport a patient to the ICU after hyperbaric oxygenotherapy	1, n=8	4/4	42 [29-46]	42 [41-44]	21 [20-23]	31 [9-39]
Tracheal tube obstruction occurring during an intra-hospital transport	You transport an intubated patient with pneumonia to the CT-scan	1, n=8	4/4	60 [51-65]	42 [38-45]	17 [17-24]	53 [24-62]
Compressive pneumothorax occurring during an intra-hospital transport	You transport an intubated patient with chest trauma to the CT-scan	1, n=8	5/3	68 [63-74]	48 [43-52]	26 [21-27]	53 [30-74]
Comparison test				Kruskal Wallis, <i>P</i> =0.0046	Kruskal Wallis, <i>P</i> <0.0001	Kruskal Wallis, <i>P</i> <0.0005	Kruskal Wallis, <i>P</i> =0.35

PGY: Post Graduate Year, TLX: Task Load IndeX, S-VAS: Visual Analogue Scale for Stress, PACU: Post Anaesthesia Care Unit, ICU: Intensive Care Unit, CT: Computered Tomography

**Table 1:** Description of scenarios, residents involved in, NASA-Task Load IndeX, skills performance and stress level across high fidelity simulation scenarios. Values are median [IQR range].

## Figures

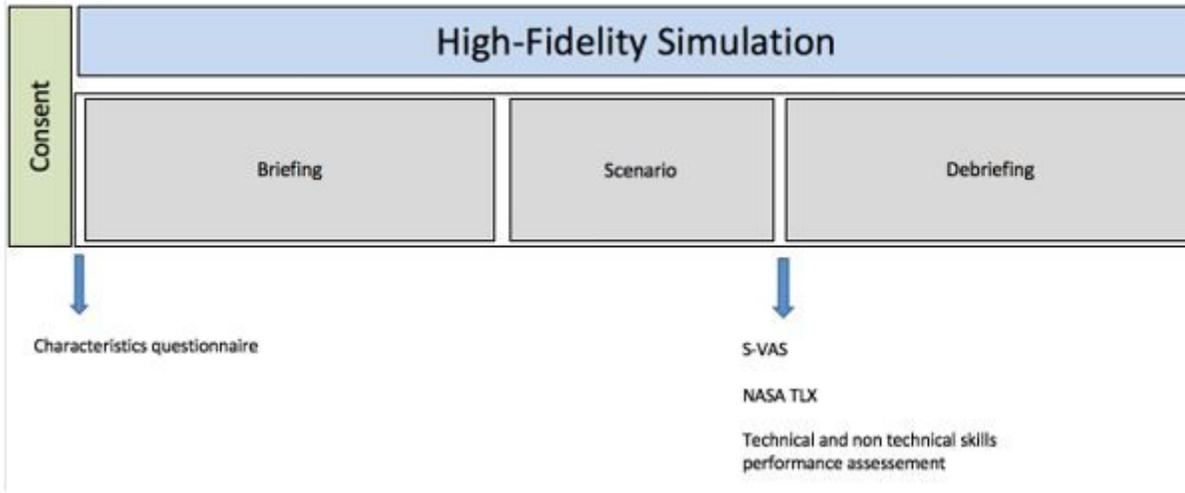


Figure 1

Timeline during high-fidelity simulation S-VAS: Visual Analogue Scale for Stress, TLX: Task Load Index

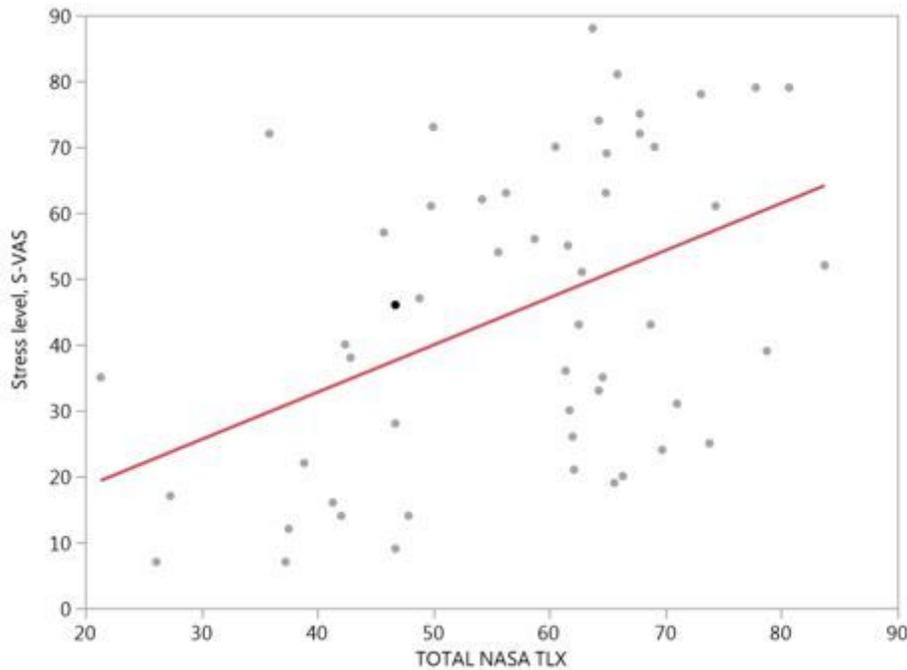


Figure 2

Bivariate fit of stress level and NASA Task Load Index score Predicted Total NASA TLX =  $44.685 + 0.279 \times \text{S-VAS}$ ,  $p=0.001$ . S-VAS: Visual Analogue Scale for Stress, TLX: Task Load Index

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplfilestitleandauthorsdetails.docx](#)

- [AppendixAHFSandNASATLXBMCMedicalEducation.docx](#)
- [AppendixBHFSandNASATLXBMCMedicalEducation.docx](#)
- [Supplfilestitleandauthorsdetails.docx](#)