

## NASA Task Load Index Scale to Evaluate the Cognitive Workload during Cardiac Anesthesia Based Simulation Scenarios

Research Article

Bhandary SP<sup>1\*</sup>, Lipps J<sup>1</sup>, Winfield SR<sup>2</sup>, Abdel-Rasoul M<sup>3</sup>, Stoicea N<sup>1</sup>, Pappada SM<sup>1</sup>, Papadimos TJ<sup>1</sup>

<sup>1</sup> Department of Anesthesiology, The Ohio State University Wexner Medical Center, Columbus, USA.

<sup>2</sup> The Clinical Skills Education and Assessment Center, The Ohio State University College of Medicine, Columbus, USA.

<sup>3</sup> Department of Statistics, The Ohio State University, Columbus, USA.

### Abstract

Simulation-based education is an important tool for anesthesiology educators given the work hour restrictions and limited exposure of anesthesiology residents to high acuity cardiac cases. Cognitive load is key to learning, performance, and resilience. When working with new and complicated information in an environment prone to changes and distractions, working memory has a limited capacity, duration, and is diminished by excessive cognitive workload. When this capacity is surpassed, learning and performance are impaired. We created cardiac anesthesia-based high fidelity simulation scenarios to determine how to best measure cognitive load by administering the NASA Task Load Index (NASA-TLX). We recreated a cardiac operating room by mimicking a complex learning environment that places high demand on the novice learners. Fourteen anesthesiology residents participated in this study for determination of cognitive load using NASA-TLX. This scale measures effort, frustration, performance, mental demand, physical demand, and temporal demand. The residents perceived mental demand as the most challenging, with a mean of  $15.21 \pm 1.86$ ; followed by effort, demand with a mean of  $14.32 \pm 2.49$ . We determined that using the NASA-TLX Scale to measure and report cognitive load provides an opportunity to effectively evaluate cognitive distress during the acquisition of new skills and enhance physician resilience.

**Keywords:** Simulation-Based Education; Cardiac Anesthesia; Cognitive Load Theory; NASA TLX Scale.

### Introduction

High fidelity simulation (HFS) provides a unique opportunity for the teaching and evaluation of knowledge, clinical reasoning, decision-making, and teamwork in a realistic and safe environment. HFS is able to uniquely address some educational needs that other educational modalities such as traditional didactics, clinical case based teaching in the operating rooms are unable to meet. In the high-acuity cardiac anesthesia setting for example, teaching residents while simultaneously caring for an anesthetized patient can be very challenging. Effective communication with surgeons, perfusionists and nursing staff is of prime importance in ensuring a successful outcome. The practice of cardiac anesthesiology also requires a distinct clinical and procedural skill set: interpretation of transesophageal echocardiography and complex hemodynamic monitoring, management of mechanical circulatory support

devices such as intra-aortic balloon pumps and ventricular assist devices, and use of multiple pharmacological measures to support the heart. With so many new technical and non-technical skills to acquire, we have found that it is easy for our residents to get overwhelmed. HFS provides an excellent educational environment to introduce some of these difficult topics without compromising patient safety. While providing learning through various clinical scenarios, instructors should always be attuned to an individual's cognitive load, as well as that of the team. Measurement of cognitive workload is a crucial aspect to the future of medical training, team interaction and professionalism.

We have developed a series of novel cardiac-based simulation scenarios meant to review difficult learning objectives of our cardiac anesthesiology rotation, and have implemented it as a part of our residency program curriculum. We hypothesized that after

#### \*Corresponding Author:

Sujatha P. Bhandary M.D.,  
Department of Anesthesiology, Division of Cardiothoracic and Vascular Anesthesiology, The Ohio State University Wexner Medical Center, Doan Hall N-411, 410 W 10th Avenue,  
Columbus, OH, 43210, USA.  
Tel: 614-293-8487  
Fax: 614-293-8153  
E-mail: [Sujatha.Bhandary@osumc.edu](mailto:Sujatha.Bhandary@osumc.edu)

**Received:** June 30, 2016

**Accepted:** July 27, 2016

**Published:** August 04, 2016

**Citation:** Bhandary SP, Lipps J, Winfield SR, Abdel-Rasoul M, Stoicea N, et al., (2016) NASA Task Load Index Scale to Evaluate the Cognitive Workload during Cardiac Anesthesia Based Simulation Scenarios. *Int J Anesth Res.* 4(8), 300-304. doi: <http://dx.doi.org/10.19070/2332-2780-1600063>

**Copyright:** Bhandary SP<sup>©</sup> 2016. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution and reproduction in any medium, provided the original author and source are credited.

completing a rotation in cardiac anesthesiology, residents will perceive these simulations to be mentally challenging, i.e., creating an increased cognitive load on the subjects. The measurement of this cognitive load, through self-reporting, may aid instructors in not only measuring the cognitive load perceived by an individual, but may also help them design and individualize scenarios.

We used the National Aeronautics and Space Administration Task Load Index Scale (NASA-TLX Scale) to determine the cognitive workload during cardiac anesthesia-based simulation scenarios to measure effort, frustration, performance, mental demand, physical demand, and temporal demand in resident subjects. The NASA-TLX Scale was developed by Hart and Staveland in 1988 to quantify the physical and mental cost or *workload* associated with performing a given task. Since its initial creation, the scale has seen widespread use in both public and private industries to evaluate the benefits and possible interference caused by a set variable, such as a new form of technology. The NASA-TLX Scale has demonstrated very low interrater variability, due to its category weighting system which accounts for an individual's self-reported strengths and weaknesses [1]. Here we demonstrate the efficacy of the self-reported NASA-TLX scale in cardiovascular anesthesia residents during HFS.

## Material and Methods

The study was conducted at The Ohio State University's Clinical Skills Education and Assessment Center using high-fidelity simulator (CAE HPS) run by a dedicated, trained simulationist under the direct supervision of the course faculty member. This study was approved by the Institutional Review Board Committee and registered in ClinicalTrials.gov (NCT02498964). The trainees provided consent.

Residents who had completed at least one month of their cardiac anesthesia rotation were included as participants in the simulation scenarios. We created 3 scenarios lasting approximately 15 minutes each. The scenarios were created by two board-certified anesthesiologists one of which is fellowship trained in cardiac anesthesiology and both of whom are simulation faculty. The scenarios were constructed using a formal scenario development process and form. Each scenario included specific clinical information and sequence of events.

A total number of 14 anesthesiology residents, postgraduate year 3 and 4 (PGY3, PGY4) volunteered to participate in the study. Based on the inclusion/exclusion criteria, the residents had already participated in at least 10 operations involving cardiopulmonary bypass (CPB) before entering the study. All the participants had received prior extensive training in high fidelity non-cardiac anesthesia scenarios as a part of their residency training. The simulation room was set up to mimic cardiac operating rooms. The monitors available to the participants were five-lead electrocardiogram (ECG), pulse-oximetry, temperature probe, capnography, arterial line pressure, central venous pressure, pulmonary artery pressure (PAP), and cardiac output (CO). Upon request, arterial blood gas values were provided and transesophageal echocardiography (TEE) images were displayed on the monitor.

A code cart, pacer box, pacing wires, and defibrillator pads were

made available. All the medications required for separation from CPB were made available in the same concentration as in the real operating rooms. Intravenous infusion pumps, inotropes, and vasopressors were primed, attached and ready for use as needed, identical to the operative room setting.

At the beginning of the each session, residents were introduced to the team. They were familiarized with the patient's medical history, preoperative evaluation and ongoing status of surgery. Each resident was asked to manage the patient and verbalize all the observations and management options as if he or she would perform in the real operating room.

## Scenarios

The first scenario (S1) involved a patient under general anesthesia and on CPB. The surgeon (course faculty) had completed an uncomplicated triple vessel coronary artery bypass graft (CABG) on a 57-year-old male with 45 minutes of bypass time. The patient presented with a junctional bradycardia (40bpm) and MAP of 50 mmHg, temperature 35°C, and SpO<sub>2</sub> of 100%. The surgeon then announces removal of the aortic cross clamp; the resident is expected to call the attending. The participant is informed that their attending will not arrive and they need to complete weaning of the patient from CPB on their own. Removal of the cross clamp results in ventricular fibrillation, and there is an expectation of communication, rhythm recognition, and appropriate intervention; pharmacological, device management, and the ordering and interpretation of pertinent laboratory data. In this scenario, hyperkalemia needed to be addressed through treatment and pacing and pharmacological support.

The second scenario (S2) begins after debriefing the previous (S1) scenario. CPB weaning has been successful. The scenario begins with a surgical request for the participant to administer protamine. At this point the patient is hemodynamically stable (V-paced at 80bpm, BP 110/60mmHg, PAP 30/15mmHg, CO 5.2 L/min), and on an epinephrine infusion at 0.02 mcg/kg/min. The patient had received 30,000 units of heparin prior to going on CPB. The last activating clotting time was 468 seconds. The patient does not experience any hemodynamic change in response to a test dose of protamine; appropriate administration continues. After 2 the patient becomes hypotension, has oxygen desaturation, a decreased CO, and pulmonary hypertension. The participant is expected to recognize the protamine reaction, discontinue the protamine infusion, call for help, communicate with the surgeon, and provide supportive care. An echocardiogram (upon request) demonstrates a dilated right ventricle. Failure to act would result in cardiovascular collapse, requiring re-initiation of CPB. The objectives for the participant in this scenario are: to demonstrate knowledge of protamine administration, the various types of protamine reactions, recognize the signs of an acute type III protamine reaction, and execute proper clinical management.

The third scenario (S3) starts following the completion and debrief of the previous scenario (S2). The chest has just been closed. The attending surgeon leaves the room, turning the case over to the surgical fellow for skin closure. Shortly thereafter ST elevation will appear in the inferior ECG leads. When this is communicated, the surgical fellow will ask for TEE images, which includes a mid-esophageal four-chamber view revealing air bubbles in the left ventricle. A trans-gastric short axis view of the heart (if

requested) will show an akinetic inferior wall. Hypotension with a decreased CO ensues. A differential diagnosis should be formed and signs/symptoms treated while communicating findings to the surgical resident. Ultimately the resident should request to the surgeon to reopen the patient's chest. Failing to do so will result in cardiogenic shock and ventricular fibrillation. The participant is expected to form a differential for post CABG hypotension and ECG changes, such as kink in the graft, pericardial effusion versus cardiac tamponade. Along with their differential, additional objectives for this scenario include diagnosing intra-cardiac air and regional wall motion abnormalities on echocardiogram, effectively treating cardiogenic shock with pharmacologic agents and communicating the need for surgical intervention if necessary.

Each resident participated as a group in all three scenarios (S1, S2, and S3). After the completion of each scenario there was an interactive debriefing session where the residents were encouraged to give feedback and to reflect upon their performance. Each debriefing session also included a structured didactic portion reviewing each clinical situation. Participation in this research project was entirely voluntary and independent of their participation in their usual didactic sessions in the simulation laboratory. The study participants filled out the NASA-TLX Scale at the end of the session. It was explained to the residents that a refusal to complete the survey would not result in penalty or loss of entitled benefits.

### Statistical Analysis

Summary statistics (mean ± standard deviation) were calculated for each of the NASA-TLX subscale components. The summary statistics were graphically represented as box plots. Student's t-tests were used to test differences in means between subscales.

All hypothesis tests were conducted at the 5% type 1 error rate. Statistical analyses were conducted using SAS software version 9.3 (SAS Institute Inc., Cary, NC).

### Results

The mental demand of the NASA-TLX Scale was the most challenging according to the residents with a mean of 15.21 ± 1.86 (95% CI: 14.14 to 16.29). The effort demand had the second highest mean of 14.32 ± 2.49 (95% CI: 12.88 to 15.76). There was not a significant difference between the mental and effort demands of the NASA-TLX Scale (p-value = 0.0994). Both mental and effort demands were significantly higher than all other NASA-TLX categories (reference p-value table). The temporal demand (mean of 12.75 ± 2.63; 95% CI: 11.23 to 14.27) also had a significantly higher mean than frustration, physical demand, and performance, respectively. The frustration (mean of 7.21 ± 5.15; 95% CI: 4.24 to 10.19), physical (mean of 6.68 ± 4.40; 95% CI: 4.14 to 9.22), and performance (mean of 8.81 ± 4.63; 95% CI: 6.01 to 11.61) aspects were not significantly different from each other [Figure 1] [Table 1] [Table 2].

### Discussion

The aim in our research was to determine the efficacy of the NASA-TLX scale to measure cognitive load on the residents during cardiac anesthesia scenarios. In doing so we would be able to determine the effects of the scenarios on effort, frustration, performance, mental demand, physical demand, and temporal demand.

Cognitive training and deliberate practice in anesthesiology are essential given the work hour restriction and limited exposure

Figure 1. Box Plot of NASA-TLX component scores for all participants.

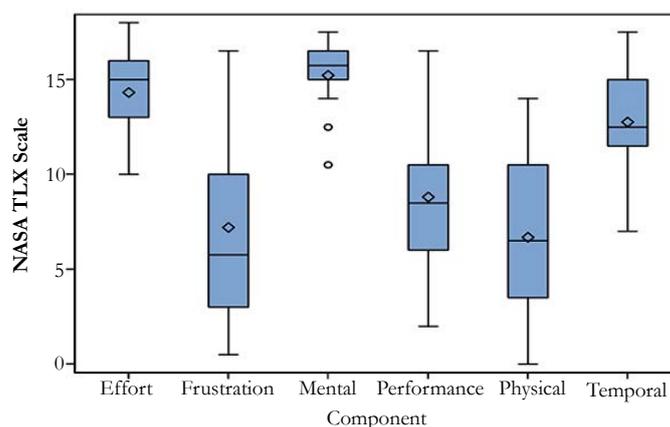


Table 1. Matrix shows the result of paired T-test comparing the different dimensions within NASA TLX. Significant differences are highlighted.

	Mental	Physical	Temporal	Performance	Effort	Frustration
Mental		<0.0001	0.0012	0.0003	0.0994	<0.0001
Physical			0.0003	0.2746	0.0001	0.7905
Temporal				0.0219	0.0162	0.0012
Performance					0.0039	0.2736
Effort						0.0001
Frustration						

**Table 2. The participants 10 PGY-4 and 4 PGY-3 residents found the scenarios to be most challenging with respect to the mental, temporal, and effort dimensions of the NASA TLX scale.**

Variable	N	Mean	Std. Dev.	Lower 95% CL	Upper 95% CL
Mental	14	15.21	1.86	14.14	16.29
Physical	14	6.68	4.4	4.14	9.22
Temporal	14	12.75	2.63	11.23	14.27
Performance	13	8.81	4.63	6.01	11.61
Effort	14	14.32	2.49	12.88	15.76
Frustration	14	7.21	5.15	4.24	10.19

to high acuity cases. Due to the high-stakes nature inherent to the cardiac operating rooms, teaching in the clinical environment can be challenging. Cardiac anesthesia cases bare similarities with crisis management. Cardiopulmonary bypass entails controlled cardiac arrests, and separation from bypass involves multitasking, resource utilization and communication with multiple team players. We created scenarios for teaching purposes that were very similar to real life cases to help residents recognize problems, respond appropriately, and reinforce the need to reassess life-threatening situations.

Cognitive load theory (CLT) assumes that the human cognitive system has a limited working memory and unlimited long-term memory. The famous “seven plus or minus two” theory assumes that working memory can hold no more than 5-9 information elements [2]. The human cognitive system is able to deal with information for no more than a few seconds. Unless refreshed by rehearsal, almost all information is lost after 20 seconds. The long-term memory holds cognitive schemas and human expertise comes from knowledge organized by these schemas. Three types of cognitive load are distinguished: intrinsic, extraneous, and germane. An intrinsic load is a direct function of the complexity of learning the task and the expertise of the learner. As in our case, our residents’ perception of cognitive load inherent to these high-stakes scenarios is based on their prior experience in similar cases in the operating room. Our residents perceived the mental, temporal and effort demands of the intrinsic workload to be challenging. The second form of cognitive load is the extraneous load that results from how the problem is presented to them. In our scenarios we attempted recreate the operating room environment so as to control the extraneous load. The third type of load is the germane load that refers to working memory resources that are used to deal with intrinsic cognitive load.

Cognitive load theory (CLT) can be used to develop instructional design principles and strategies based on a human cognitive architecture [3]. Design guidelines based on CLT principle aim to decrease extraneous load, manage intrinsic load and optimize germane load. The best method of education and training is still unknown. However, expert demonstration and participation in HFS seem to be more effective in teamwork training than traditional didactics [4]. Deliberate practice and learner centered models have been used to improve skills as there is a shift in focus from apprenticeship model to competency based model [5]. Competence is a spectrum ranging from novice, to advanced beginner, to competent, to proficient to expert [6]. The scenarios in this research project were created for the novice learner.

Cognitive task analysis (CTA) has been used in surgical education and has been shown to improve surgical outcome parameters including time, precision, accuracy and error reduction in both real world and simulated environments [7]. Simulation-based training has been shown to improve non-technical skills in surgical residents [8]. Likewise, senior anesthesiology residents that received simulation-based training in weaning from CPB have demonstrated improved performance in actual patient care as rated by their instructors [9]. Theory behind CTA has been widely adopted by various disciplines. The NASA-TLX Scale has previously been used and validated for cognitive workload assessment and training in diverse non-healthcare fields. Military officials have used this concept to master difficult skills. In addition it has been used by musicians, pilots, Olympic athletes, and engineering. While use of the NASA-TLX Scale for cognitive workload assessment is new to the health care field, it appears to provide a reliable measure of intrinsic cognitive load [10]. We believe this tool can be useful for both the trainee and the instructor in optimizing cognitive load thereby facilitating both technical and non-technical skill acquisition for residents learning these difficult tasks. In fact, the authors have been involved in the development of an artificial neural network (ANN) algorithm, through iterative modeling and the use of EEG and bioradios, to produce a remote real time NASA-TLX cognitive work load output for training. In this way each subject becomes an individual NASA-TLX model.

Our approach in creating these scenarios was to identify problems in intraoperative decision-making, and immediately follow the scenarios with a debriefing to impart knowledge in an efficient manner. After completion of the scenarios we used the NASA-TLX scale to determine the residents’ cognitive load. Use of the NASA-TLX scale allows for better design of scenarios, through determining each individual’s cognitive workload. Scenarios can be adjusted for each individual. Doing so will help teachers in medicine be aware of cognitive distress, and in realizing this aspect in an individual, may help facilitation of resilience in individuals. Future research will also involve team cognitive workload indicators, as well as measurements in professional interactions.

Simulation provides a unique opportunity to teach using multi-sensory engagement. Unlike traditional didactics, our scenarios allowed for participants to interact and perform in a physical space. While our goal was to mentally challenge the participants, we also aimed to evaluate the other dimensions of cognitive load. We demonstrated that the NASA-TLX scale can be used to assess

each of the dimensions/demands for trainees undergoing high fidelity simulations.

It is not enough to rehearse difficult critical scenarios in a safe, simulated environment. We must identify through observation and reflection, those areas in which workload is highest and address them in order to improve as physicians. Workload is a concept which has traditionally been difficult to measure-but the NASA TLX gives both learners and educators a chance to do so. Through honest self-reflection we can improve our performance and thereby patient care. Future studies might look at the utility of the scale outside of the simulation lab and in the actual ORs during routine care, better identifying areas of high workload by various members of the healthcare team and leading to process improvement.

### Limitations

Our study is limited by the fact that we are a single center looking at a homogenous group of learners at a single point in time. This study was not randomized and our subject numbers were small. We have not studied the transfer of this knowledge to resident performance in the operating rooms. Every resident received the simulation-based training as a part of the curriculum. Both PGY3 and PGY4 residents were exposed to this training, and as such, PGY4 residents would have performed more cardiac cases in the operating rooms than their junior counterparts. Even though each simulation session was followed by an interactive debriefing session with didactic teaching to reinforce the knowledge, our study does not compare different teaching methods.

As well, participants in our study were instructed to rate their work load for all three scenarios rather than separately. Furthermore, while all three participants were involved in each scenario a different individual took on the lead role. We observed scenario two and three to be more difficult than the first, yet the NASA-TLX scale was completed after the 3<sup>rd</sup> scenario which may have had an effect on the perceived cognitive workload [1].

Our study is also inherently limited by our method of evaluation, the NASA-TLX Scale. While the scale is widely considered to be the gold standard upon which other workload scales are judged, there is no way for the NASA-TLX Scale to accurately capture every workload-influencing aspect of a situation. Moreover, many of these aspects are contested within the current psychological literature [1]. NASA TLX ratings have also been shown to be affected by context effect, the influence of experiences on the rater before or after the event, and by range effect whereby raters fail to use the whole range of the scale. The range effect might be offset in future studies by providing more explicit instructions to the participants [1].

### Conclusion

The NASA-TLX Scale is instrumental in measuring the cognitive workload in anesthesiology residents practicing in a simulated environment. This tool can also help guide instructors

in developing scenarios that are appropriately challenging to residents with respect to each of the NASA-TLX scale's domains and assess the progress made by the resident. It has been used effectively in military settings throughout the U.S. armed forces. From a teacher's perspective there can be a determination of cognitive load during acquisition of knowledge in the simulated environment, and help design future scenarios for trainees at all levels of experience. As academicians we have a duty to train and develop competent, well-rounded teachers and leaders in medicine. Simulation-based training with measurement of cognitive workload provides an opportunity to safely practice and improve skills without exposing real patients to the possibility of adverse events. This is especially important in the case of high stress medical fields, procedures and interactions. Cognitive workload measurement is an important and challenging aspect of the future in medical teaching.

### Declaration of Interest

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

### References

- [1]. Hart S G (2006) Nasa-Task Load Index (NASA-TLX); 20 Years Later. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 50(9): 904-908. doi: 10.1177/154193120605000909.
- [2]. Miller G A (1956) The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological review* 63(2): 81-97. doi: 10.1037/h0043158
- [3]. Van Merriënboer J J, Sweller J (2010) Cognitive load theory in health professional education: design principles and strategies. *Medical education* 44(1): 85-93. doi: 10.1111/j.1365-2923
- [4]. Semler MW, Keriwala RD, Clune JK, Rice TW, Pugh ME, et al. (2015) A Randomized Trial Comparing Didactics, Demonstration, and Simulation for Teaching Teamwork to Medical Residents. *Annals of the American Thoracic Society*, 12(4): 512-519. doi: 10.1513/AnnalsATS.201501-030OC
- [5]. Bhatti N I, Ahmed A (2015) Improving skills development in residency using a deliberate-practice and learner-centered model. *The Laryngoscope* 125(S8): S1-S14. doi: 10.1002/lary.25434
- [6]. Benner P (1982) From novice to expert. *AJN The American Journal of Nursing*, 82(3): 402-407. doi: 10.1002/nur.4770080119
- [7]. Wingfield LR, Kulendran M, Chow A, Nehme J, Purkayastha S (2014) Cognitive Task Analysis Bringing Olympic Athlete Style Training to Surgical Education. *Surgical innovation* 22(4): 406-17. doi: 10.1177/1553350614556364
- [8]. Nguyen N, Elliott JO, Watson WD, Dominguez E (2015) Simulation Improves Nontechnical Skills Performance of Residents During the Perioperative and Intraoperative Phases of Surgery. *Journal of surgical education* 72(5): 957-63. doi: 10.1016/j.jsurg.2015.03.005
- [9]. Bruppacher HR, Alam SK, LeBlanc VR, Latter D, Naik VN (2010) Simulation-based training improves physicians' performance in patient care in high-stakes clinical setting of cardiac surgery. *Anesthesiology*, 112(4): 985-992. doi: 10.1097/ALN.0b013e3181d3e31c
- [10]. Naismith LM, Cheung JJ, Ringsted C, Cavalcanti R B (2015) Limitations of subjective cognitive load measures in simulation-based procedural training. *Medical education* 49(8): 805-814. doi: 10.1111/medu.12732